DISCOVERY AND CHEMICAL OPTIMIZATION OF NOVEL RHO KINASE (ROCK II) AND/OR Ca^{2+}/CALMODULIN-DEPENDENT PROTEIN KINASE II (CaMKII) INHIBITORS AS POTENTIAL NEW ANTI-HYPERTENSIVE AGENTS.

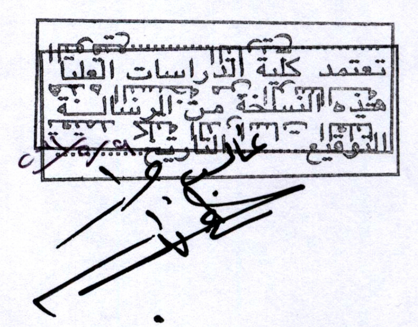
By
Rand Omar Mohammad Shahin

Supervisor
Dr. Mutasem O. Taha, Prof.

This Thesis was submitted in Partial Fulfillment of the Requirements for the Doctor of Philosophy Degree in Pharmaceutical Sciences/Medicinal Chemistry and Drug Discovery Faculty of Graduate Studies

The University of Jordan

May, 2011



This Thesis / Dissertation (Discovery and chemical optimization of novel Rho Kinase (ROCK II) and/or Ca^{2+}/Calmodulin-Dependent protein Kinase II (CaMKII) as potential new anti-hypertensive agents.) was successfully defended and approved on 7th April 2011.

Examination Committee Signature

Dr. Mutasem Taha, Supervisor
Prof. of Medicinal Chemistry and Drug Design ---------------------

Dr. Mohammad S. Mubarak
Prof. of Organic Chemistry ---------------------

Dr. Imad I. Hamdan
Prof. of Pharmaceutical Analysis ---------------------

Dr. Mohammad M. Hudaib
Associate Prof. of Phytochemistry and Phytochemical Analysis ---------------------

Dr. Amjad M. Qandil
Associate Prof. of Organic and Medicinal Chemistry
(Jordan University of Science and Technology) ---------------------

Dedicated

To my Family

Father, Mother

My Beloved Husband

Daughters Yusra and Salma

My sisters Raghad and Saja

Who offered me unconditional love and

Support throughout the course of this thesis,

Without your support, this would have never

Been possible.

Acknowledgments

There are a number of people without whom this thesis might

not have been

written, and to whom I am greatly indebted.

My deepest thanks go to my father Prof. Omar Shahin,

whose dedication inspirited me to learn the importance of

scarification as an essential tool for success.

I must also thank my supervisor who shepherded me

through the

bulk of the work, Prof. Mutasem Taha, for being there from

start to finish,

providing me with his valuable guidance and advice. I am

especially

grateful to him for constantly giving me the motivation to

perform to my

maximum ability. It has been a rewarding learning

experience to work

under his guidance.

Thanks due to the members of the examination committee;

Prof.

Mohammad Mubarak, Prof. Imad Hamdan, Dr.

Mohammad Hudaib & Dr.

Amjad Qandil for their suggestions & comments, and for

devoting some of

their valuable time to read and evaluate this work.

I am also grateful for the help and expertise provided to me

by the members of the mecicinal chemistry lab.; Saja

Qutaishat , Heba el-egil and Nadia Hijaz for their

support, words of encouragement and the tremendous

amount of goodwill.

Finally, I am forever indebted to husband and mother for

their endless patience

and encouragement when it was most required. Without the

stability and

security provided by their love, this study would not have

been possible.

List of Contents

List of Tables

List of Figures

List of Schemes

Abbreviation	Definition
^{13}C NMR	Carbon-13 nuclear magnetic resonance
^{1}H NMR	Proton nuclear magnetic resonance
3D	Three dimensional
3D-QSAR	Three-dimensional Structure-Activity Relationship
ACC	Overall Accuracy
AUC	Area Under the Curve
CaMKII	Ca2+/Calmodulin-dependent Protein Kinase II
CRD	Cysteine-Rich repeat Domain
CVD	CardioVascular Diseases
DMSO	Dimethyl sulfoxide
DS 2.0	Discovery Studio, Version 2.0
FNR	Overall false negative rate
GA	Genetic Algorithm
GFA	Genetic Function Approximation
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
Hbic	Hydrophobic
Hypo	Hypotheses
MLR	Multiple Linear Regression
NCI	National Cancer Institute
NegIon	Negative Ionizable
NMR	Nuclear Magnetic Resonance
NP	Count of Non-polar atoms (i.e. Cl, F, Br, I, S and C atoms)
NumHBA	The Number of Hydrogen-Bond Acceptors
NumHBD	The Number of Hydrogen-Bond Donors
PDB	Protein Data Bank
PH	Pleckstrin-Homology domain
QSAR	Quantitative Structure Activity Relationship
$r^2_{5\text{-fold}}$	Cross-validation correlation coefficient determined by the leave-20%-
r^2_{LOO}	Cross-validation correlation coefficient determined by the leave-one-
RBD	Rho-Binding Domain
RingArom	Ring Aromatic
RMSD	Root of the Mean of Squared differences
ROC	Receiver Operating Characteristic
ROCK II	Rho Associated protein Kinases type II
Se	Sensitivity
Sp	Specificity
SPC	Overall specificity
TPR	Overall true positive rate

DISCOVERY AND CHEMICAL OPTIMIZATION OF NOVEL RHO KINASE (ROCK II) AND/OR $Ca^{2+/}$CALMODULIN-DEPENDENT PROTEIN KINASE II (CaMKII) AS POTENTIAL NEW ANTI-HYPERTENSIVE AGENTS

By

Rand Omar Shahin
Supervisor
Dr. Mutasem O. Taha, Prof.

Abstract

The pharmacophoric space of both Rho Kinase (ROCK II) and $Ca^{2+/}$Calmodulin-dependent Protein Kinase II a (CaMKII) were explored using several diverse sets of known ROCK II and CaMKII inhibitors. Subsequently, genetic algorithm and multiple linear regression analyses were employed to select an optimal combination of pharmacophoric models and physicochemical descriptors that access self-consistent quantitative structure-activity relationship (QSAR) of optimal prediction power for ROCK II (r = 0.82, F-statistic = 18.75, r^2_{LOO} = 0.68, r^2_{PRESS} = 0.54), Two orthogonal pharmacophores (of cross-correlation r^2 = 0.28) emerged in the QSAR equation suggesting the existence of at least two distinct binding modes accessible to ligands within ROCK II binding pocket as for CaMKII (r^2 = 0.70, F-statistic = 18.19, r^2_{LOO} = 0.71, r^2_{PRESS} = 0.60). Three orthogonal pharmacophores (of cross-correlation $r^2 \leq 0.21$) emerged in the QSAR equation suggesting the existence of at least three distinct binding modes accessible to ligands within CaMKII binding pocket. The resulted pharmacophores were further validated via receiver-operating characteristic (ROC) curve profiles. The validity of QSAR equation and its pharmacophores was experimentally established by *in vitro* assay. Eight NCI hits exert low micromolar ROCK II inhibition, while two NCI hits exert nanomolar CaMKII inhibition and many other exert low micromolar inhibition. Chemical modifications were carried out based on one of the most active NCI hits. Triazine derivatives were synthesized and tested as anti-CaMKII compounds with IC_{50} = 154 nM for the most active one and the rest are of various inhibition percentages.

Chapter One

Elaborate Ligand-Based Modeling and Subsequent Synthetic Optimization Access New Nanomolar Ca^{2+}/Calmodulin-Dependent Protein Kinase II Inhibitors

1.1. Chapter abstract

Ca^{2+}/Calmodulin-Dependent Protein Kinase II (CaMKII) has been recently implicated in cardiovascular diseases and hypertension prompting several attempts to discover and optimize new CaMKIIδ inhibitors. Towards this end we explored the pharmacophoric space of 88 CaMKIIδ inhibitors using nine diverse sets of inhibitors to identify high quality pharmacophores. Subsequently, genetic algorithm and multiple linear regression analysis were employed to select an optimal combination of pharmacophoric models and 2D physicochemical descriptors capable of accessing self-consistent quantitative structure-activity relationship (QSAR) of optimal predictive potential (r^2_{72} = 0.70, F = 18.19, r^2_{LOO} = 0.71, r^2_{PRESS} against 16 external test inhibitors = 0.60). Three orthogonal pharmacophores emerged in the QSAR equation suggesting the existence of at least three binding modes accessible to ligands within CaMKIIδ binding pocket. Receiver operating characteristic (ROC) curves analysis established the validity of QSAR-selected pharmacophores. We employed the pharmacophoric models and associated QSAR equation to screen the national cancer institute (NCI) list of compounds. *In silico* screening identified nanomolar and low micromolar inhibitors. The most potent hits exhibited IC_{50} values of 20 and 82 nM. QSAR equation and pharmacophoric models were employed to guide the synthesis of novel triazine-based CaMKIIδ inhibitors, of which the most potent illustrated an IC_{50} value of 154 nM against CaMKIIδ.

.

1.2. Introduction

Ca^{2+}/calmodulin-dependent protein kinase II (CaMKII), which is a member of Ca^{2+}/calmodulin-dependent protein kinases (three subtypes I, II and IV), is a multifunctional serine/threonine kinase that acts as an ubiquitous mediator of Ca^{2+}-linked signaling. It phosphorylates a wide range of substrates involved in regulating Ca^{2+} mediated alterations in cellular function (Hudmon & Schulman, 2002).

There are four CaMKII isoforms (□, □, □, and □) distributed in various tissues: CaMKII□ and □ are found primarily in the brain, while CaMKII□ is found in the heart (Hund *et al*, 2008). CaMKII□ is involved in diverse myocardial roles including regulation of excitation, contraction, transcription, and apoptosis. Therefore, it represents an attractive target for therapeutic intervention in cardiovascular diseases (Hund *et al*, 2008; Anderson *et al*, 2005; Sossalla *et al*, 2010; Stolen *et al*, 2009; MacDonnell *et al*, 2009; Kohlhaas *et al*, 2006)

Like other Ca^{2+}/calmodulin-dependent protein kinases, CaMKII□ is regulated by calmodulin (CaM), which is an ubiquitous Ca^{2+} sensor protein that activates substrates proteins and enzymes in a Ca^{2+}-dependent manner. CaM binding to CaMKII□ induces conformational changes in CaMKII□ that relieves autoinhibition and allows full enzyme activity (Anderson *et al*, 2005; Kurokawa *et al*, 2001)

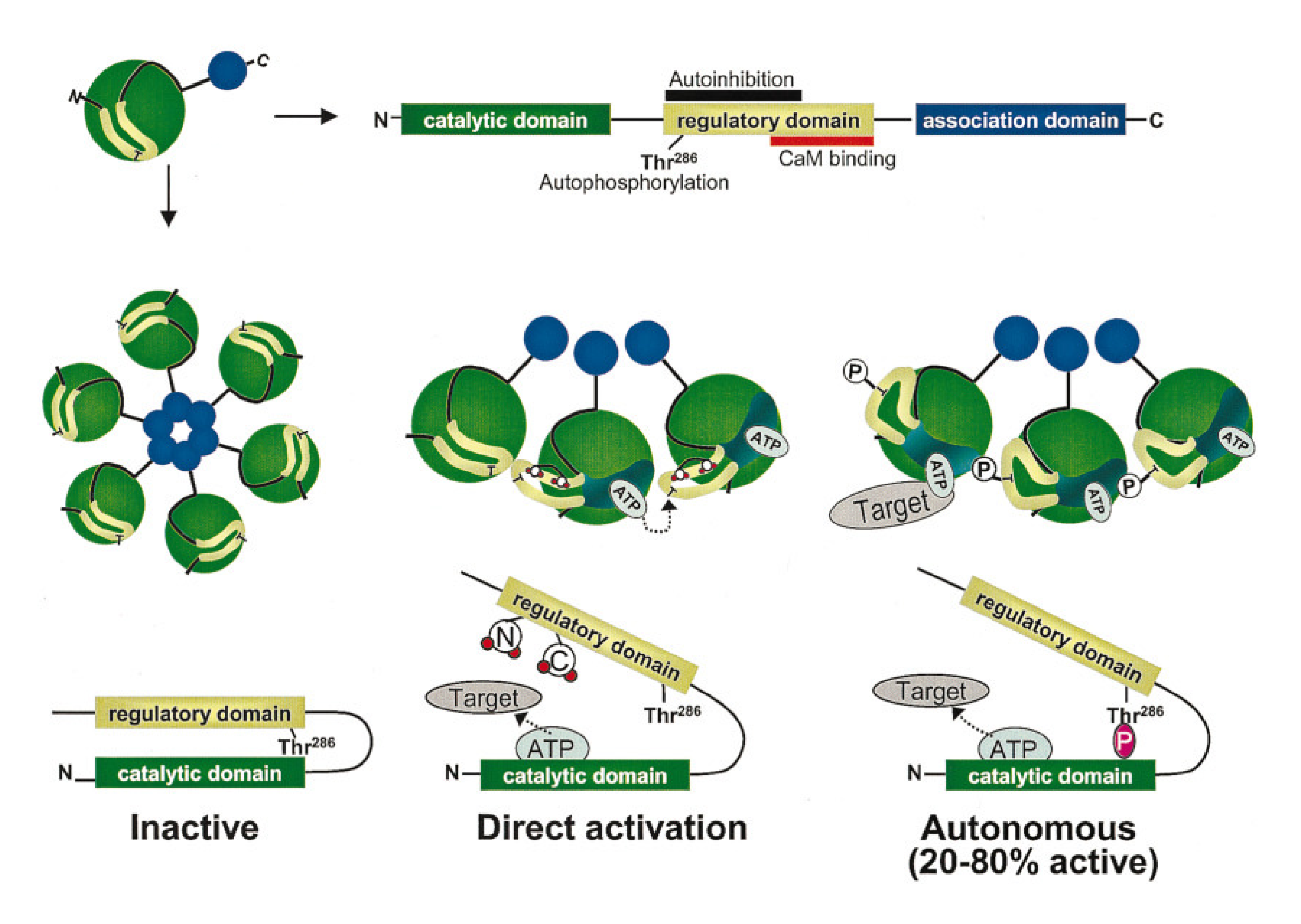


Figure 1.2 Domain layout and oligomeric organization of CaMKII. The three main domains of the CaMKII monomer are indicated in cartoon and linear layout (top). Middle left shows that CaMKII forms homo- or heteromultimers (6-12 monomers) in wheel-like structures (a second one may sit on top of the one shown, forming a double wheel).Lower middle and right panels show activation of CaMKII by Ca-CaM binding and subsequent autophosphorylation at Thr286 (P).CaM binding is sufficient to activate CaMKII so the active site (ATP) can interact and phosphorylation target proteins, but autophosphorylation makes CaMKII active (20-80%) even after CaM dissociates (Maier & Bers, 2002).

The validity of CaMKII as a target for therapeutic intervention in cardiovascular diseases was established by emerging clinical trials employing potent CaMKII inhibitors: KN-62, KN-93 and the natural CaMKII inhibitor staurosporine (Hudmon & Schulman, 2002).

However, despite their high cellular activity and clinical progression, CaMKII□ inhibitors suffer from several limitations, such as, poor cellular permeability, general toxicity, and specificity (Hudmon & Schulman, 2002). These issues have led to significant efforts to identify novel small molecular inhibitors of CaMKII□ (Hudmon & Schulman, 2002, Levy *et al*, a & b , 2008 , Lu *et al*, 2008 , Mavunkel *et al*, 2008).

The unavailability of satisfactory high resolution crystallographic structures for CaMKII□ confined most discovery projects to high throughput screening and homology modeling, i.e., molecular modeling studies that utilize known structures of other kinases to generate predictive structural information about CaMKII (Hudmon & Schulman, 2002).

The continued interest in designing new CaMKII□ inhibitors combined the unavailability of CaMKII□ X-ray complexes and lack of adequate ligand-based computer-aided drug discovery efforts prompted us to explore the possibility of developing ligand-based three-dimensional (3D) pharmacophore(s) integrated within self-consistent QSAR model. The pharmacophore model(s) can be used as 3D search queries to discover new CaMKII□ inhibitory scaffolds, while the QSAR model can be used to predict the bioactivities of captured hits and to prioritize them for *in vitro* testing. We previously reported the use of this innovative approach towards the discovery of new inhibitory leads against glycogen synthase kinase-3β (Taha *et al*, 2008 , a,) , bacterial MurF (Taha *et al,* 2008 , b), protein tyrosine phosphatase (Taha *et al,* 2007), DPP IV (Al-masri *et al,* 2008), hormone sensitive lipase (Taha *et al*, 2008, c), □-secretase (Al-Nadaf *et al*, 2010), influenza neuraminidase (Abu-Hammad *et al,* 2009), cholesteryl ester transfer protein (Abu Khalaf *et al*, 2010), cycline dependent kinase (Al-Sha'er *et al*, 2010 , a), Heat shock protein (Al-Sha'er *et al*, 2010 , b), estrogen receptor β (Taha *et al,* 2010), □-D-Glucosidase (Abu

Khalaf *et al*, DOI 10.1007/s00894-010-0737-1), and β-D-Galactosidase (Abdula *et al*, 2011).

We employed the CATALYST-HYPOGEN module of the software package Discovery Studio (Discovery Studio 2.5.5 User Guide., 2010) to construct plausible binding hypotheses for a diverse list of CaMKIIδ inhibitors (Levy *et al*, a & b , 2008 , Lu *et al*, 2008 , Mavunkel *et al*, 2008). Subsequently, genetic function algorithm (GFA) and multiple linear regression (MLR) analyses were employed to search for an optimal QSAR that combine high-quality binding pharmacophores with other molecular descriptors and capable of explaining bioactivity variation across a collection of diverse CaMKIIδ inhibitors. The optimal pharmacophores were subsequently used as 3D search queries to screen the national cancer institute (NCI) list of compounds for new CaMKIIδ inhibitory leads that were tested *in vitro*. The resulting active hits and pharmacophores were used to guide synthesis of novel nanomolar triazine-based CaMKII inhibitors.

1.3. Experimental

1.3.1 Molecular Modeling

1.3.1.1 Software and Hardware

The following software packages were utilized in the present research.

- CATALYST (Version 4.11), Accelrys Inc. (www.accelrys.com), USA.
- CERIUS2 (Version 4.10), Accelrys Inc. (www.accelrys.com), USA.
- CS ChemDraw Ultra 6.0, Cambridge Soft Corp. (http:// www.cambridgesoft.Com), USA.
- Discovery Studio 2.5, Accelrys Inc. (www.accelrys.com), USA.

Pharmacophore and QSAR modeling studies were performed using CATALYST (HYPOGEN module) and CERIUS2 software suites from Accelrys Inc. (San Diego, California, www.accelrys.com) installed on a Silicon Graphics Octane2 desktop workstation equipped with a dual 600 MHz MIPS R14000 processor (1.0 GB RAM) running the Irix 6.5 operating system. Structure drawing was performed employing ChemDraw Ultra 6.0 which was installed on a Pentium 4 PC.

1.3.1.2 Pharmacophore Modelling

1.3.1.2.1 Data Set

The structures of 88 CaMKIIδ inhibitors (**1**-**88**, Table 1.1) were collected from recently published literature (Levy *et al.,* 2008, a&b; Lu *et al.,* 2008; Mavunkel *et al.,* 2008). Although the *in vitro* bioactivities of the collected inhibitors were gathered from four separate articles they were determined employing the same bioassay methodology that

involves incubation of the CaM kinase II sample with substrate, in the presence of $Mg^{2+}$and ^{32}P-labeled ATP. The bioactivities were expressed as the concentrations of the test compounds that inhibited the activity of CaMKII by 50% (IC_{50}, μM). The logarithm of measured IC_{50} (μM) values were used in the three-dimensional quantitative structure activity analysis (3D-QSAR), thus correlating the data linear to the free energy change.

In cases where IC_{50} is expressed as being higher than 20 μM (e.g., compounds **4**, **8**, **14**, **29**, **79** and **80**, see Table 1.1) it was assumed it equals 210 μM. This assumption is necessary to allow 3.5 log cycles separation between the most potent training compound (i.e., **85**) and the least active training compounds (i.e., **4**, **8**, **14**, **29**, **79** and **80**). This is necessary for pharmacophore modeling in CATALYST. Moreover, this assumption is necessary to allow statistical correlation and QSAR analysis since regression requires exact values. The logarithmic transformation of IC_{50} values should minimize any potential errors resulting from this assumption.

The two-dimensional (2D) chemical structures of the inhibitors were sketched using ChemDraw Ultra and saved in MDL-molfile format. Subsequently, they were imported into CATALYST, converted into corresponding standard 3D structures and energy minimized to the closest local minimum using the molecular mechanics CHARMm force field implemented in CATALYST. The resulting 3D structures were utilized as starting conformers for CATALYST conformational analysis.

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

O, R_1, R_3, N, O, X, N, R_2

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
1	1	H N, *	NH_2, *	H	---	CH	0.38
2	1	H N, *	O, HN, O, *	H	---	CH	3.81
3	1	S, *	NH_2, *	H	---	CH	0.36
4[**]	1	S, *	O, HN, O, *	H	---	CH	>20
5	1	*	NH_2, *	H	---	CH	0.29

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

No.	Reference [a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
6	1		NH₂	H	---	CH	0.5
7	1	N	NH₂	H	---	CH	0.63
8	1	N	NH₂	H	---	CH	>20
9	1		NH₂	H	---	CH	1.35
10	1	N	NH₂	H	---	CH	1.58
11**	1	N	NH₂	H	---	CH	0.54

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

O, R_1, R_3–N, O, X, N–R_2

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
12	1	N, *	NH_2, *	H	---	CH	1.6
13[**]	1	O, *	NH_2, *	H	---	CH	1.49
14	1	S, *	O, HN, O, *	H	---	CH	>20
15**	1	S, *	NH_2, *	H	---	CH	1.19

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
16	1		NH_2	H	---	CH	0.28
17	1		NH_2	H	---	CH	1.32
18	1	O	NH_2	H	---	CH	1.68
19	1	N	NH_2	H	---	CH	4.56

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
20	1			H	---	CH	0.75
21[**]	1			H	---	CH	7.61
24	1			CH_3	---	CH	15.02
22	1			H	---	CH	0.034
23	1			CH_3	---	CH	2.32

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

No.	Reference[a]	R$_1$	R$_2$	R$_3$	Y	X	IC$_{50}$ (µM)
24	1	*	NH$_2$ *	CH$_3$	---	CH	15.02
25	1	H N Br *	NH$_2$ *	CH$_3$	---	CH	>20
26	2	H N *	H	H	---	CH	1.19
27	2	H N Br *	CH$_3$	H	---	CH	0.87
28	2	*	CH$_3$	H	---	CH	>20

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
29	2		H	H	---	CH	>20
30	2		H		---	CH	0.18
31[**]	2		H	H	---	CH	3.53
32	2			H	---	CH	7.54

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
33	2			H	---	CH	>20
34	2			H	---	CH	2.02
35	2			H	---	CH	1.45
36	2			H	---	CH	1.89
37	2			H	---	CH	0.31

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
38	2	[structure]	[structure]	H	---	CH	0.64
39[**]	2	[structure]	[structure]	H	---	CH	1.87
40	2	[structure]	[structure]	H	---	CH	0.45
41	3	[structure]	[structure]	H	---	CCH_3	0.74
42	3	[structure]	[structure]	H	---	N	0.38

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
43	3	NH_2	H	---	---	---	0.19
44	3	N	H	---	---	---	0.74
45	3	OH	H	---	---	---	6
46	3	NH_2	H	---	---	---	0.11
47[**]	3	H N	H	---	---	---	0.15
48	3	N	H	---	---	---	0.56
49	3	N	H	---	---	---	2.41

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
50[**]	3	* NH	H	---	---	---	0.68
51	3	* NH_2	Cl	---	---	---	0.01
52	3	* H N	Cl	---	---	---	0.013
53	3	* NH_2	F	---	---	---	0.038
54	3	* NH_2	CF_3	---	---	---	0.032
55[**]	3	* NH_2	CN	---	---	---	0.046
56	3	* H N	CH_2CH_3	---	---	---	0.11
57	3	* NH_2	OCH3	---	---	---	0.17

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

O, H, N, O, N, R1, R2, N, H

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
58[**]	3	H N *	CO_2CH_3	---	---	---	0.021
59[**]	3	H N *	O OH *	---	---	---	0.38
60	3	H N *	O NH_2 *	---	---	---	0.19
61	3	H N *	O NH *	---	---	---	0.28
62	3	H N *	O N *	---	---	---	11.85

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

Y N R$_1$ N H N X

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (µM)
63	4	* NH	---	---	N	F *	1.28
64	4	* NH	---	---	N	*	3.14
65	4	* NH	---	---	N	*	0.92
66	4	* NH	---	---	N	Cl *	2.87
67	4	* NH	---	---	N	O OCH_3 *	1.69
68	4	* NH	---	---	N	O OH *	0.21

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

Y N N R1 N H X

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
69	4	* NH	---	---	N	HO O *	12.99
70	4	* NH	---	---	N	* O HO	3.42
71	4	* NH	---	---	N	* N	2.22
72	4	* NH	---	---	N	* O HN	0.72
73	4	* NH	---	---	N	* H N N N N	2.26
74	4	* NH	---	---	N	* O N—H HO	0.17
75	4	NH_2	---	---	N	* O OH	0.066
76	4	H	---	---	N	* O OH	0.39

Table 1.1: The structures of CaMKδII inhibitors utilized in modeling.

Y N R1 N H N X

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
76	4	H	---	---	N	* O OH	0.39
77	4	H	---	---	CH	* O OH	0.064
78	4	Cl	---	---	CH	* O OH	0.042
79**	4	H	---	---	CH	* O HO	>20
80	4	H	---	---	CH	* O OCH_3	>20
81	4	H	---	---	CH	* O HN—	>20
82**	4	H	---	---	CH	* O N—H HO	0.29
83**	4	H	---	---	CH	* H N N N N	0.77

Table1.1: The structures of CaMKδII inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Y	X	IC_{50} (μM)
84	4	Cl	---	---	CH	NH2	0.66
85	4	Cl	---	---	CH	N NH	0.009
86	4	Cl	---	---	CH	N O	0.24
87[**]	4	Cl	---	---	CH	N	0.93
88	4	Cl	---	---	CH	NH NH2	0.012

[a]References: 1: Levy *et al*, a , 2008 ; 2: Levy *et al*, b , 2008, 3: Lu *et al*, 2008 , 4 : Mavunkel *et al*, 2008.
* Connection point
** These compounds were employed as external testing set for QSAR validation

1.3.1.2.2 Conformational Analysis

The molecular flexibilities of the collected compounds were taken into account by considering each compound as a collection of conformers representing different areas of the conformational space accessible to the molecule within a given energy range. Accordingly, the conformational space of each inhibitor (**1-88**, Table 1.1) was explored adopting the "best conformer generation" option within CATALYST (CATALYST 4.11 users' manual, 2005) based on the generalized CHARMm force field implemented in the program. Default parameters were employed in the conformation generation procedure of training compounds, i.e., a conformational ensemble was generated with an energy threshold of 20 kcal/mol from the local minimized structure at which has the lowest energy level and a maximum limit of 250 conformers per molecule (CATALYST 4.11 users' manual, 2005).

1.3.1.2.3 Pharmacophoric Hypotheses Generation

All 88 molecules with their associated conformational models were regrouped into a spreadsheet. The biological data of the inhibitors were reported with an "Uncertainty" value of 3, which means that the actual bioactivity of a particular inhibitor is assumed to be situated somewhere in an interval ranging from one-third to three-times the reported bioactivity value of that inhibitor (Kurogi and Güner, 2001; Sutter *et al.*, 2000; Poptodorov *et al.*, 2006). Subsequently, nine structurally diverse training subsets: sets **I**, **II**, **III**, **IV**, **V**, **VI, VII, VIII and IX** in table 1.2, respectively were carefully selected from the collection for pharmacophore modeling. Typically, CATALYST requires informative training sets that include at least 16 compounds of

evenly spread bioactivities over at least three and a half logarithmic cycles. Lesser training lists could lead to chance correlation and thus faulty models.

Table 1.2: Training subsets employed in exploring the pharmacophoric space of CaMK□II inhibitors.

subsets	Most active[a]	Moderate active[a]	Least active[a]	No. of compounds
I	22, 51, 53, 54, 85, 88	3, 6, 9, 15, 16, 24, 30, 31, 37, 42, 43, 46, 68, 86	28, 79, 80	23
II	53, 54, 55, 58, 85, 88	3, 6, 9, 15, 16, 22, 23, 24, 26, 27, 31, 37, 42, 57.	28, 29, 79, 80, 81	25
III	22, 51, 53, 54, 85, 88	6, 9, 15, 24, 31, 68, 86	25, 28, 79	16
IV	53, 54, 55, 58, 85, 88	3, 6, 9, 15, 16, 22, 23, 24, 26, 27, 31, 37, 42, 57.	28, 29, 79, 80, 81	25
V	57, 77, 87	23, 31, 64, 66, 67, 68, 69, 70, 73, 78, 82,83	25, 28, 29, 80, 81	20
VI	53, 54, 55, 58, 78, 85, 88	3,5,6,9,15,23,24,26,27, 31,37,48,57,61,82	28, 29, 79, 80, 81	27
VII	53, 54, 55, 58, 78, 85, 88	3,5,6,9,15,23,24,26,27, 31,37,48,57,61 82	28, 29, 79, 80, 81	27
VIII	22, 51, 53, 54, 78, 88	3, 6, 9, 15, 16, 24, 26, 30, 31, 37, 43, 46, 66, 67, 68, 70, 87, 82, 86	28, 79, 80, 81	29
IX	53, 54, 55, 58	3, 6, 9, 15, 16, 22, 23, 24, 26, 27, 31, 37, 42, 57	28, 29, 79, 80, 81	23

[a]Potency categories as defined by equations (1.1) and (1.2).

The selected training sets were utilized to conduct 68 modeling runs to explore the pharmacophoric space of CaMKIIδ inhibitors (table 1.3). The exploration process included altering interfeature spacing parameter (100 and 300 picometers) and the maximum number of allowed features in the resulting pharmacophore hypotheses, i.e., they were allowed to vary from 4 to 5 for first and second runs and from 5 to 5 for third and fourth runs of each training set, as shown in table 1.3.

Pharmacophore modeling employing CATALYST proceeds through three successive phases: the constructive phase, subtractive phase and optimization phase During the constructive phase, CATALYST generates common conformational alignments among the most-active training compounds. Only molecular alignments based on a maximum of five chemical features are considered. The program identifies a particular compound as being within the most active category if it satisfies equation (1.1) (Poptodorov *et al.*, 2006; Kurogi and Güner, 2001; Li *et al.*, 2000).

$$(\text{MAct} \times \text{UncMAct}) - (\text{Act} / \text{UncAct}) > 0.0 \quad \ldots\ldots\ldots\ldots\ldots (1.1)$$

Where "MAct" is the activity of the most active compound in the training set, "Unc" is the uncertainty of the compounds and "Act" is the activity of the training compounds under question (Table 1.2).

In the subsequent subtractive phase, CATALYST eliminates some hypotheses that fit inactive training compounds. A particular training compound is defined as being inactive if it satisfies equation (1.2) (Poptodorov *et al.*, 2006; Kurogi and Güner, 2001; Li *et al.*, 2000):

$$\text{Log}(\text{Act}) - \log(\text{MAct}) > 3.5 \quad \ldots\ldots\ldots\ldots (1.2)$$

Among the inactives, some inhibitors were reported to have IC_{50} > 20 µM (e.g., **4, 8, 14, 29, 79 and 80**). These were fed into CATALYST with IC_{50} values = 210 µM. This assumption should have minimal impact on the pharmacophore modeling process as inactive compounds are used to filter out poor pharmacophores regardless to their actual bioactivities.

However, in the optimization phase, CATALYST applies fine perturbations in the form of vectored feature rotation, adding new feature and/or removing a feature, to selected hypotheses that survived the substractive phase, in an attempt to find new models of enhanced bioactivity/mapping correlations. CATALYST selects the highest-ranking models (10 by default) and presents them as the optimal pharmacophore hypotheses resulting from the particular automatic modeling run.

Eventually, our pharmacophore exploration efforts (68 automatic runs, Tables 1.2 and 1.3) culminated in 680 pharmacophore models of variable qualities.

Table 1.3: Training sets and CATALYST run parameters employed for exploring CaMKδII pharmacophoric space

Run Number	Training set[a]	Number of compounds	Selected features[b]	Max-Min Features[c]	Spacing[d]
1	I	23	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	100
2		23	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	300
3		23	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	100
4		23	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	300
5		23	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1) , ExVol (0-10)	4-5	100
6		23	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	4-5	300
7		23	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	5-5	100
8		23	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	5-5	300
9	II	25	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	100
10		25	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	300
11		25	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	100
12		25	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	300
13		25	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1) , ExVol (0-10)	4-5	100
14		25	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	4-5	300
15		25	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	5-5	100
16		25	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	5-5	300
17	III	16	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	100
18		16	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	300
19		16	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	100
20		16	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	300
21		16	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1) , ExVol (0-10)	4-5	100
22		16	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	4-5	300
23		16	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	5-5	100
24		16	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	5-5	300
25	IV	25	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	100
26		25	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	300
27		25	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	100
28		25	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	300
29		25	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1) , ExVol (0-10)	4-5	100
30		25	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	4-5	300
31		25	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	5-5	100
32		25	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1), ExVol (0-10)	5-5	300

Table 1.3: Training sets and CATALYST run parameters employed for exploring CaMKδII pharmacophoric space

Run Number	Training set[a]	Number of compounds	Selected features[b]	Max-Min Features[c]	Spacing[d]
33	V	20	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), NegIon (0-1)	4-5	100
34		20	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), NegIon (0-1)	4-5	300
35		20	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), NegIon (0-1)	5-5	100
36		20	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), NegIon (0-1)	5-5	300
37		20	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), NegIon (0-1) , ExVol (0-10)	4-5	100
38		20	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), NegIon (0-1), ExVol (0-10)	4-5	300
39		20	HBA (0-5), HBD (0-5), Hbic (0-5), ExVol (0-10)	5-5	100
40		20	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), NegIon (0-1), ExVol (0-10)	5-5	300
41	VI	29	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3)	4-5	100
42		29	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3)	4-5	300
43		29	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3)	5-5	100
44		29	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3)	5-5	300
45		29	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	4-5	100
46		29	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	4-5	300
47		29	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	5-5	100
48		29	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	5-5	300
49	VII	27	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3)	4-5	100
50		27	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3)	4-5	300
51		27	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3)	5-5	100
52		27	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3)	5-5	300
53		27	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	4-5	100
54		27	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	4-5	300
55		27	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	5-5	100
56		27	Hbic(0-3), HBA (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	5-5	300
57	VIII	27	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3)	4-5	100
58		27	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3)	4-5	300
59		27	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3)	5-5	100
60		27	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3)	5-5	300
61		27	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	4-5	100
62		27	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	5-5	100
63		27	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	4-5	300
64		27	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), ExVol (0-10)	5-5	300

Table 1.3: Training sets and CATALYST run parameters employed for exploring CaMKδII pharmacophoric space

Run Number	Training set[a]	Number of compounds	Selected features[b]	Max-Min Features[c]	Spacing[d]
65	VII	23	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	100
66		23	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	4-5	300
67		23	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	100
68		23	Hbic(0-3), HBAF (0-3), HBD (0-3), RingArom (0-3), PosIon (0-1)	5-5	300

[a]Training subsets as in table 1.2.
[b]HBA: Hydrogen Bond Acceptor, HBD: Hydrogen Bond Donor, HBAF: Hydrogen Bond Acceptors including fluorine atoms, RingArom: Ring Aromatic, Hbic: Hydrophobic, HbicArom : Hydrophobic Aromatic, PosIon: Positive Ionizable, NegIon: Negative Ionizable, ExVol: Exclusion Volumes. Numbers in brackets refer to the allowed ranges of corresponding features in each pharmacophore modeling run.
[c]Max-Min refers to the allowed range of pharmacophoric features in each model.
[d]Spacing refers to the maximum interfeature distance in picometers. Other parameters were set to their default values.

1.3.1.2.4 Assessment of the Generated Hypotheses

When generating hypotheses, CATALYST attempts to minimize a cost function consisting of three terms: Weight cost, Error cost and Configuration cost (CATALYST 4.11 users' manual, 2005; Poptodorov *et al*., 2006; Kurogi and Güner, 2001; Li *et al*., 2000; Sutter *et al*. 2000; Bersuker *et al*. 2000). Weight cost is a value that increases as the feature weight in a model deviates from an ideal value of 2. The deviation between the estimated activities of the training set and their experimentally determined values adds to the error cost. The activity of any compound can be estimated from a particular hypothesis through equation (1.3) (CATALYST 4.11 users' manual, 2005).

$$\text{Log (Estimated Activity)} = \text{I+Fit} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \; (1.3)$$

Where, I = the intercept of the regression line obtained by plotting the log of the biological activity of the training set compounds against the Fit values of the training compounds. The Fit value for any compound is obtained automatically employing equation (1.4) (CATALYST 4.11 users' manual, 2005).

$$\text{Fit} = \Sigma \text{ mapped hypothesis features} \times W\,[1 - \Sigma\,(\text{disp/tol})^2] \ldots\ldots\ldots \; (1.4)$$

Where, Σ mapped hypothesis features represents the number of pharmacophore features that successfully superimpose (i.e., map or overlap with) corresponding chemical moieties within the fitted compound, W is the weight of the corresponding hypothesis feature spheres. This value is fixed to 1.0 in CATALYST-generated models. disp is the distance between the center of a particular pharmacophoric sphere (feature centroid) and the center of the corresponding superimposed chemical moiety of the fitted compound; tol is the radius of the pharmacophoric feature sphere (known as Tolerance, equals to 1.6 Å by default). $\Sigma\,(\text{disp/tol})^2$ is the summation of $(\text{disp/tol})^2$ values for all pharmacophoric features that successfully superimpose corresponding

chemical functionalities in the fitted compound. (CATALYST 4.11 users' manual, 2005).

The third term, i.e., the configuration cost, penalizes the complexity of the hypothesis. This is a fixed cost, which is equal to the entropy of the hypothesis space. The more the numbers of features (a maximum of five) in a generated hypothesis, the higher is the entropy with subsequent increase in this cost. The overall cost (total cost) of a hypothesis is calculated by summing over the three cost factors. However, error cost is the main contributor to total cost.

CATALYST also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are normally distributed about their mean. Accordingly, the greater the difference from the null hypothesis cost, the more likely that the hypothesis does not reflect a chance correlation. In a successful automatic modeling run, CATALYST ranks the generated models according to their total costs (CATALYST 4.11 users' manual, 2005).

An additional approach to assess the quality of CATALYST-HYPOGEN pharmacophores is to cross-validate them using the Cat-Scramble program implemented in CATALYST. This validation procedure is based on Fisher's randomization test (Fisher *et al*, 1966). In this validation test, a 95% confidence level was selected, which instruct CATALYST to generate 19 random spreadsheets by the Cat-Scramble command. Subsequently, CATALYST-HYPOGEN is challenged to use these random spreadsheets to generate hypotheses using exactly the same features and parameters used in generating the initial unscrambled hypotheses (Krovat *et al*. 2003). Success in generating pharmacophores of comparable cost criteria to those produced by the original unscrambled data reduces the confidence in the training compounds and the unscrambled original pharmacophore models. Based on Fischer

randomization criteria; only 677 pharmacophores exceeded the 85% significance threshold for subsequent processing (clustering and QSAR analyses).

1.3.1.2.5 Clustering of the Generated Pharmacophore Hypotheses

The successful models (677) were clustered into 134 groups utilizing the hierarchical average linkage method available in CATALYST. Subsequently, the highest-ranking representatives, as judged based on their significance F-values, were selected to represent their corresponding clusters in subsequent QSAR modeling. Table 1.4 shows information about representative pharmacophores including their pharmacophoric features, success criteria and differences from corresponding null hypotheses. The table also shows the corresponding Cat.

1.3.1.3 QSAR Modeling

A subset of 72 compounds from the total list of inhibitors (**1-88**, Table 1.1) was utilized as a training set for QSAR modeling; the remaining 16 molecules (*ca*. 20% of the dataset) were employed as an external test subset for validating the QSAR models. The test molecules were selected as follows: the 88 inhibitors were ranked according to their IC_{50} values, and then every fifth compound was selected for the test set starting from the high-potency end. This selection considers the fact that the test molecules must represent a range of biological activities similar to that of the training set. The selected test inhibitors are marked with double asterisks in table 1.1.

The logarithm of measured $1/IC_{50}$ (µM) values was used in QSAR, thus correlating the data linear to the free energy change. The chemical structures of the inhibitors were imported into CERIUS2 as standard 3D single conformer representations in SD

format. Subsequently, different descriptor groups were calculated for each compound employing the C2.DESCRIPTOR module of CERIUS2. The calculated descriptors included various simple and valence connectivity indices, electro-topological state indices and other molecular descriptors (e.g., logarithm of partition coefficient, polarizability, dipole moment, molecular volume, molecular weight, molecular surface area, etc.) (CERIUS2 QSAR Users' Manual, 2005). The training compounds were fitted (using the best-fit option in CATALYST (CATALYST 4.11 users' manual, 2005) against the representative pharmacophores (134 models, table 1.4), and their fit values were added as additional descriptors. The fit value for any compound is obtained automatically via equation (1.4) (CATALYST 4.11 users' manual, 2005).

Genetic function approximation (GFA) was employed to search for the best possible QSAR regression equation capable of correlating the variations in biological activities of the training compounds with variations in the generated descriptors, i.e., multiple linear regression modeling (MLR). GFA techniques rely on the evolutionary operations of ''crossover and mutation'' to select optimal combinations of descriptors (i.e., chromosomes) capable of explaining bioactivity variation among training compounds from a large pool of possible descriptor combinations, i.e., chromosomes population. However, to avoid overwhelming GFA-MLR with large number of poor descriptor populations, we removed lowest-variance descriptors (20%) prior to QSAR analysis.

Each chromosome is associated with a fitness value that reflects how good it is compared to other solutions. The fitness function employed herein is based on Friedman's 'lack-of-fit' (LOF) (CERIUS2 QSAR Users' Manual, 2005)..

Our preliminary diagnostic trials suggested the following optimal GFA parameters: explore linear, quadratic and spline equations at mating and mutation probabilities of

50%; population size = 500; number of genetic iterations = 30,000 and lack-of-fit (LOF) smoothness parameter = 1.0. However, to determine the optimal number of explanatory terms (QSAR descriptors), it was decided to scan and evaluate all possible QSAR models resulting from 5 to 25 explanatory terms.

All QSAR models were validated employing leave one-out cross-validation (r^2_{LOO}), bootstrapping (r^2_{BS}) and predictive r^2 (r^2_{PRESS}) calculated from the test subsets. The predictive r^2_{PRESS} is defined as:

$$r^2_{PRESS} = SD\text{-}PRESS/SD \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1.5)$$

Where SD is the sum of the squared deviations between the biological activities of the test set and the mean activity of the training set molecules, PRESS is the squared deviations between predicted and actual activity values for every molecule in the test set.

1.3.1.4 Addition of exclusion volumes

account for the steric constrains of the binding pocket we decided to decorate Hypo9/47 and Hypo7/39 with exclusion volumes employing Hip-Hop-Refine module of CATALYST. Hip-Hop-Refine uses inactive training compounds to construct excluded volumes that resemble the steric constrains of the binding pocket. It identifies spaces occupied by the conformations of inactive compounds and free from active ones. These regions are then filled with excluded volumes (Al-masri, *et al.*, 2008; Taha, *et al.*, 2005; Taha, *et al.*, 2007; Taha, *et al.*, 2008; Al-Sha'er and Taha., 2010; CATALYST 4.11 users' manual, 2005).

A subset of 23 compounds (table 1.7) was used for constructing appropriate exclusion regions around Hypo9/47 and Hypo7/39.

In Hip-Hop-Refine the user defines how many molecules must map completely or partially the particular hypothesis via the Principal and Maximum Omitted Features (MaxOmitFeat) parameters. Active compounds are normally assigned MaxOmitFeat parameter of zero and principal value of 2 to instruct the software to consider all their chemical moieties in pharmacophore modeling and to fit them against all the pharmacophoric features of a particular hypothesis. On the other hand, inactive compounds are allowed to miss one (or two) features by assigning them a MaxOmitFeat of 1 (or 2) and principal value of zero.

To construct the Hip-Hop refine list we decided to consider 210 μM as an appropriate activity/inactivity threshold. Accordingly, inhibitors of IC_{50} values < 210 μM were regarded as ‘‘actives’’ and were assigned principal and MaxOmitFeat values of 2 and 0, respectively, On the other hand, inhibitors of IC_{50} > 210 μM were considered inactive and were assigned principal values of zero (CATALYST 4.11 users’ manual, 2005).

On the other hand, inhibitors of IC_{50} values ranging from 0.080 to 210 μM were considered as intermediates and were assigned Principal values of 1 and MaxOmitFeat parameter of 1 (rarely 0 or 2). While inhibitors of IC_{50} values > 210 μM were regarded as inactives and were assigned a Principal value of 0. However, each inactive compound was carefully evaluated to assess whether its low potency is attributable to missing one or more pharmacophoric features, i.e., compared to active compounds, or related to possible steric clashes within the binding pocket, or due to both factors. Therefore, inactive compounds suspected of missing one or more pharmacophoric features were assigned MaxOmitFeat values of 1 or 2, respectively. Spaces occupied by conformers and/or mappings of this group of compounds and free

from conformers and/or mappings of active compounds are filled with excluded volumes. However, compounds that seem to be inactive mainly due to steric clashes within the binding pocket were assigned MaxOmitFeat value of zero. This value instructs HipHop-Refine to force inactive compound(s) to map all the pharmacophoric features of the binding model, and therefore permits the software to identify spaces occupied by excess structural fragments/features of such inactive compounds and fill them with excluded volumes (Taha *et al.*, 2007). HipHop-Refine was configured to allow a maximum of 100 exclusion spheres to be added to Hypo9/47 and Hypo7/39. Table 1.7, shows the training compounds employed in this step and their corresponding principal and MaxOmitFeat parameters.

1.3.1.5 Receiver Operating Characteristic (ROC) Curve Analysis

Selected pharmacophore models (i.e., Hypo8/31, Hypo7/39 and Hypo 9/47) were validated by assessing their abilities to selectively capture diverse CaMKII active compounds from a large testing list of actives and decoys.

$$\ldots\ldots(1.6)\; D(i,j) = \sqrt{(NumHBD_i - NumHBD_j)^2 + (NumHBA_i - NumHBA_j)^2 + (NP_i - NP_j)^2}$$

The testing list was prepared as described by Verdonk and co-worker (Kirchmair *et al*, 2008; Verdonk *et al*, 2004). Briefly, decoy compounds were selected based on three basic one-dimensional (1D) properties that allow the assessment of distance (D) between two molecules (e.g., i and j): (1) the number of hydrogen-bond donors (NumHBD); (2) number of hydrogen-bond acceptors (NumHBA) and (3) count of nonpolar atoms (NP, defined as the summation of Cl, F, Br, I, S and C atoms in a

particular molecule). For each active compound in the test set, the distance to the nearest other active compound is assessed by their Euclidean Distance (Eq. (1.6)):

The minimum distances are then averaged over all active compounds (Dmin). Subsequently, for each active compound in the test set, around 30 decoys were randomly chosen from the ZINC database (Irwin and Shoichet, 2005). The decoys were selected in such a way that they did not exceed Dmin distance from their corresponding active compound.

To diversify active members in the list, we excluded any active compound having zero distance ($D(i, j)$) from other active compound(s) in the test set. Active testing compounds were defined as those possessing CaMKII affinities ranging from 0.009 to 15.02 μM. The test set included 39 active compounds and 1307 ZINC decoys.

The test set (1346 compounds) was screened by each particular pharmacophore employing the "Best flexible search" option implemented in CATALYST, while the conformational spaces of the compounds were generated employing the "Fast conformation generation option" implemented in CATALYST. Compounds missing one or more features were discarded from the hit list. *In-silico* hits were scored employing their fit values as calculated by Eq. (1.4).

The ROC curve analysis describes the sensitivity (Se or true positive rate, Eq. (1.7)) for any possible change in the number of selected compounds (n) as a function of (1-Sp). Sp is defined as specificity or true negative rate (Eq. (1.8)) (Kirchmair *et al*., 2008; Triballeau *et al*., 2005).

$$Se = \frac{Number\ of\ Selected\ Actives}{Total\ Number\ of\ Actives} = \frac{TP}{TP + FN} \quad \text{..............................} \quad (1.7)$$

$$Sp = \frac{Number\ of\ Discarded\ Inactives}{Total\ Number\ of\ Inactives} = \frac{TN}{TN + FP} \quad \text{........ ....} \quad (1.8)$$

where, TP is the number of active compounds captured by the virtual screening method (true positives), FN is the number of active compounds discarded by the virtual screening method, TN is the number of discarded decoys (presumably inactives), while FP is the number of captured decoys (presumably inactive) (Kirchmair *et al*., 2008; Triballeau *et al*., 2005).

If all molecules scored by a virtual screening (VS) protocol with sufficient discriminatory power are ranked according to their score (i.e., fit values), starting with the best-scored molecule and ending with the molecule that got the lowest score, most of the actives will have a higher score than the decoys. Since some of the actives will be scored lower than decoys, an overlap between the distribution of active molecules and decoys will occur, which will lead to the prediction of false positives and false negatives (Kirchmair *et al*., 2008; Triballeau *et al*., 2005). The selection of one score value as a threshold strongly influences the ratio of actives to decoys and therefore the validation of a VS method. The ROC curve method avoids the selection of a threshold by considering all Se and Sp pairs for each score threshold (Kirchmair *et al*., 2008; Triballeau *et al*., 2005). A ROC curve is plotted by setting the score of the active molecule as the first threshold. Afterwards, the number of decoys within this cutoff is counted and the corresponding Se and Sp pair is calculated. This calculation is repeated for the active molecule with the second highest score and so forth, until the scores of all actives are considered as selection thresholds.

The ROC curve representing ideal distributions, where no overlap between the scores of active molecules and decoys exists, proceeds from the origin to the upper-left corner until all the actives are retrieved and Se reaches the value of 1. In contrast to that, the ROC curve for a set of actives and decoys with randomly distributed scores tends towards the Se = 1-Sp line asymptotically with increasing number of actives and decoys (Kirchmair *et al*., 2008; Triballeau *et al*., 2005). The success of a particular virtual screening workflow can be judged from the following criteria:

1) Area under the ROC curve (AUC) (Triballeau *et al*., 2005). In an optimal ROC curve an AUC value of 1 is obtained; however, random distributions cause an AUC value of 0.5. Virtual screening that performs better than a random discrimination of actives and decoys retrieve an AUC value between 0.5 and 1, whereas an AUC value lower than 0.5 represents the unfavorable case of a virtual screening method that has a higher probability to assign the best scores to decoys than to actives (Kirchmair *et al*., 2008; Triballeau *et al*., 2005).
2) Overall accuracy (ACC): describes the percentage of correctly classified molecules by the screening protocol (equation 1.9). Testing compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) (Gao *et al*., 1999; Jacobsson *et al*., 2003; Kirchmair *et al*., 2008).

$$ACC = \frac{TP + TN}{N} = \frac{A}{N}.Se + \left(1 - \frac{A}{N}\right).Sp \quad \text{.................(1.9)}$$

where, N is the total number of compounds in the testing database, A is the number of true actives in the testing database.

3) Overall specificity (SPC): describes the percentage of discarded inactives by the particular virtual screening workflow. Inactive test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured)

regardless to their individual fit values (Gao *et al*., 1999; Jacobsson *et al*., 2003; Kirchmair *et al*., 2008).

4) Overall true positive rate (TPR or overall sensitivity): describes the fraction percentage of captured actives from the total number of actives. Active test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values.
5) Overall false negative rate (FNR or overall percentage of discarded actives): describes the fraction percentage of active compounds discarded by the virtual screening method. Discarded active test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values.

1.3.1.6 *In Silico* Screening for New CaMKIIδ Inhibitors

Hypo8/31 and the sterically refined versions of Hypo9/47 and Hypo7/39 were employed as 3D search queries to screen the NCI 3D flexible structural database. The screening was done employing the "Best Flexible Database Search" option implemented within CATALYST. NCI hits were filtered according to Lipinski's and Veber's rules (Lipinski *et al.,* 2001; Veber *et al*., 2002), as in table (1.9). Remaining hits were fitted against the three pharmacophores using the "best fit" option within CATALYST. The fit values together with the relevant molecular descriptors of each hit were substituted in the optimal QSAR equation (1.10). The highest ranking molecules based on QSAR predictions were acquired and tested *in vitro*. Table 1.10 shows active hits and their QSAR-predictions and experimental bioactivities.

1.3.2 Synthesis of Triazine Analogues

General Synthetic Procedures

Melting points were measured using Gallenkampf melting point apparatus and are uncorrected. ^{1}H NMR and ^{13}C NMR spectrums were collected on a Bruker Avance Ultrashield NMR400 MHz spectrometer. High resolution mass spectrometry was performed using LC Mass Bruker Apex-IV mass spectrometer utilizing an electrospray interface. Infrared spectra were recorded using Shimadzu IR Affinity-1 spectrophotometer. The samples were analyzed as KBr pellets. Analytical thin layer chromatography (TLC) was carried out using pre-coated aluminum plates and visualized by UV light (at 254 and/or 360 nm). Elemental analysis was performed using EuroVector elemental analyzer. Chemicals and solvents were used without further purification.

1.3.2.1 Preparation of 4,6-Dichloro-1,3,5-triazin-2-ylamino derivatives (99–103, Scheme 1.1)

To a magnetically-stirred, ice-bathed, solution of substituted aniline (1 equivalent) in acetone (20 mL) at 0 °C was added cyanuric chloride (1 equivalent), and the resulting mixture was stirred at 0°C for 2 h and then at room temperature for an additional 2h. Subsequently, crushed ice (10 mL) was added to the reaction, and the mixture was allowed to warm up to room temperature over 1 h. The solid was collected by vacuum filtration and was washed with water 3 times (3-10 mL) (Leftheris *et al*, 2004).

(4,6-Dichloro-[1,3,5]triazin-2-yl)-(2-methoxy-4-methyl-phenyl)-amine (99)

Prepared from commercially available 4-methyl o-anisidine (3.321 gm, 27 mmol) and cyanuric chloride (5 gm, 27 mmol) to yield **99** as cream colored powder (3.4 gm, 44%). mp: 176 °C from acetone, ν_{max} (KBr disc) 3381, 1724, 1620, 1579 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 2.24 (s, 3H, CH_3), 3.73 (s, 3H, OCH_3), 6.98 (m,1H), 7.07 (m,1H), 7.14 (s,1H), and 10.46 (s, 1H, NH) ppm;^{13}C NMR (100 MHz, DMSO-d_6): δ 22.70 (CH_3), 58.42 (OCH_3), 114.88 (CH), 126.77 (C), 129.84 (CH), 131.07 (CH), 131.91 (C), 154.09 (C), 167.59 (C), 171.50 and 172.24 (C, rotamer splitting) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{11}H_{11}Cl_2N_4O$: 285.03044 found: 285.03040. Anal. Calcd for $C_{11}H_{10}Cl_2N_4O$: C, 46.34; H, 3.54, N, 19.65 found: C, 46.39, H, 3.37, N, 19.29.

(4,6-Dichloro-[1,3,5]triazin-2-yl)-(2-methoxy-phenyl)-amine (100)

Prepared from commercially available *o*-anisidine (0.74 gm, 5.4 mmol) and cyanuric chloride (1 gm, 5.4 mmol) to yield **100** as cream colored powder (0.75 gm, 51%). mp: 171 °C from acetone (lit. mp: 175 °C from CCl_4) (Desai *et al*, 2004); ν_{max} (KBr disc) 3393, 1608, 1570 cm^{-1}, ^{1}H NMR (400 MHz, DMSO-d_6): δ 3.78 (s, 3H, OCH_3), 6.96 (m, 1H), 7.12 (m, 1H), 7.26 (m, 1H), 7.39 (m, 1H), and 10.48 (s, 1H, NH) ppm;^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 58.37 (OCH_3), 114.88 (CH), 122.97 (CH), 126.95 (C), 129.35 (CH), 130.75 (CH), 156 (C), 167.49 (C), 171.54 and 172.25 (C, rotamer splitting) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{10}H_9Cl_2N_4O$: 271.01479 found: 271.01470. Anal. Calcd for $C_{10}H_8Cl_2N_4O$: C, 44.30; H, 2.97, N, 20.67 found: C, 42.81, H, 3.07, N, 19.57.

(5-Chloro-2-methyl-phenyl)-(4,6-dichloro-[1,3,5]triazin-2-yl)-amine (101)

Prepared from commercially available 5-Chloro-2-methyl-aniline (0.78 gm, 5.4 mmol) and cyanuric chloride (1.0 gm, 5.4 mmol) to yield **101** as rose colored powder (1.16 gm, 74%). mp: 192.5 °C from acetone (lit. mp: 192-193) (Fontana *et al*, 1960); ν_{max} (KBr disc) 3225, 1608, 1587 cm^{-1} , ^{1}H NMR (400 MHz, DMSO-d_6): δ 2.23 (s, 3H, CH_3), 7.22 (m, 2H), 7.66 (s, 1H), 11.34 (s, 1H, NH) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{10}H_8Cl_3N_4$: 288.98091 found: 288.98086. Anal. Calcd for $C_{10}H_7Cl_3N_4$: C, 41.48; H, 2.44, N, 19.35. Found: C, 41.10, H, 2.53, N, 18.54.

(4-Chloro-phenyl)-(4,6-dichloro-[1,3,5]triazin-2-yl)-amine (102)

Prepared from commercially available 4-chloro-aniline (2.06 gm, 16.2 mmol) and cyanuric chloride (3.0 gm, 16.2 mmol) to yield **102** as white colored powder (3.4 gm, 76%). mp: 189-190 °C from acetone (lit. mp: 187-188 from benzene) (Curd *et al,*1947); ν_{max} (KBr disc) 3396, 1768, 1722, 1601, 1541 cm^{-1}, ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.39 (m, 2H), 7.57 (m, 2H), 10.96 (br s, 1H, NH) ppm. ^{13}C NMR (100 MHz, DMSO-d_6,): δ 125.50 (2xCH), 131.14 (2xCH), 138.47 (C), 156.55 (C), 166.31 (C), 171.45 and 172.34 (C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_9H_6Cl_3N_4$: 274.96526 found 274.96469. Anal. Calcd for $C_9H_5Cl_3N_4$: C, 39.23; H, 1.83, N, 20.33 found: C, 39.64, H, 1.75, N, 20.11.

(4,6-Dichloro-[1,3,5]triazin-2-yl)-(4-nitro-phenyl)-amine (103)

Prepared from commercially available 4-nitro-aniline (3.8 gm, 27.5 mmol) and cyanuric chloride (5 gm, 27.3 mmol) to yield **103** as yellow powder (7.5 gm, 95%). mp: 288 °C from acetone (lit. mp: 290 from ethanol) (Desai *et al,* 1990); ν_{max} (KBr

disc) 3344, 1618, 1575, 1539 cm^{-1}, , ^{1}H NMR (400 MHz, DMSO-d_6): ▯ 7.70 (m, 2H), 8.18 (m, 2H), and 10.94 (br s, 1H, NH) ppm. HRMS-ESI *m/z* [*M*-H]$^+$ calcd for $C_9H_4Cl_2N_5O_2$: 283.97475 found: 283.97475.

1.3.2.2 Preparation of 6-chloro-1,3,5-triazin-2,4-diamine derivatives (104–113, scheme 1.2)

To a solution of the particular 4,6-dichloro-[1,3,5]triazin-2-yl)amine (**99-103**) in THF (50 ml) at 0°C, diisopropylethylamine (acid scavenger, 2 equivalent) and the selected alkylamine (2 equivalent) were added sequentially. The resulting mixture was stirred at 0 °C for 30 min, then at room temperature for another 30 min before quenching with aqueous solution of ammonium chloride (50% w/v). Water/THF were totally removed in vacuo and the product was washed with water and purified by recrystallization from acetone/water (Leftheris *et al*, 2004).

6-Chloro-N-ethyl-N'-(2-methoxy-phenyl)-[1,3,5]triazine-2,4-diamine (104)

Prepared from **100** (2.25 gm, 8.3 mmol) and ethylamine (0.75 gm, 16.6 mmol) to yield **104** as cream-colored powder (1.84 g, 79%), mp: 178°C from THF; ν_{max} (KBr disc) 3387, 3261, 1626, 1575, 1518 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): ▯ 1.10 (t, 3H), 3.2 (q, 2H), 3.80 (s, 3H, OCH_3), 6.91-7.1 (m, 4H), 7.3 (br d, 1H, NH) and 8.15(br s, 1H, NH) ppm. HRMS-ESI *m/z* [*M*+H]$^+$ calcd for $C_{12}H_{15}Cl_1N_5O$: 280.09596 found: 280.09544. Anal. Calcd for $C_{12}H_{14}Cl_1N_5O$: C, 51.52; H, 5.04, N, 25.04 found: C, 51.23, H, 4.83, N, 24.45.

6-Chloro-N-isopropyl-N'-(2-methoxy-phenyl)-[1,3,5]triazine-2,4-diamine (105)

Prepared from **100** (2 gm, 7.3 mmol) and isopropylamine (0.86 gm, 14.6 mmol) to yield **105** as off-white colored powder (4.5 gm, 100%). mp: 100 °C from THF; HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{13}H_{17}ClN_5O$: 294.11161 found: 294.11148.

6-Chloro-N-(2-methoxy-4-methyl-phenyl)-N'-pentyl-[1,3,5]triazine-2,4-diamine (106)

prepared from **99** (2 gm, 7 mmol) and pentylamine (1.2 gm , 14 mmol) to yield **106** as white colored powder (2.3 gm, 98%), mp 159-165 °C from THF, ^{1}H NMR (400 MHz, DMSO-d_6 , rotameric (Leftheris *et al*, 2004)): δ 0.81 (m, 3H, CH_3), 1.24 (m, 4H, 2xCH_2), 1.44 (m, 2H, CH_2), 2.23 (s, 3H, CH_3), 3.18 (m, 2H, CH_2), 3.75 (s, 3H, OCH_3), 6.88-6.93 (m, 3H), 7.53 (brs, exchangeable, NH), and 8.1 (brs, exchangeable, NH) ppm. ^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 16.54 (CH_3), 23.04 (CH_2), 24.44 (CH_2), 30.89 (CH_2), 31.11 (CH_2), 31.23 (CH_3), 58.43 (OCH_3), 113.78 (CH), 114.09 (CH), 127.73 (C), 128.85 (C), 131.43 (CH), 152.20 (C), 166.53 (C), 167.92 (C), and 170.52 (C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{16}H_{23}ClN_5O$: 336.15856 found: 336.15867; Anal. Calcd for $C_{16}H_{22}ClN_5O$: C, 57.22; H, 6.60, N, 20.85. found: C, 56.49, H 5.94, N, 20.11.

Chloro-N-isopropyl-N'-(2-methoxy-4-methyl-phenyl)-[1,3,5]triazine-2,4-diamine (107)

prepared from **99** (1 gm, 3.5 mmol) and isopropylamine (0.41 gm, 7.0 mmol). to yield **107** as white powder (0.7, 65%). mp: 145 °C from THF; ν_{max} (KBr disc) 3396, 3256, 1581, 1493 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6 , rotameric (Leftheris *et al*, 2004)): δ 1.20 (d, 6H, 2xCH_3), 2.21 (s, 3H, CH_3), 3.04 (m, 1H), 3.71(s, 3H, OMe), 6.83 (m, 3H), 7.64 (br s,1H, NH), and 7.95 (br s,1H, NH) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{14}H_{19}ClN_5O$: 308.12726 found: 308.12731.

6-Chloro-N-(5-chloro-2-methyl-phenyl)-N'-ethyl-[1,3,5]triazine-2,4-diamine (108)

Prepared from **101** (1.0 gm, 3.45 mmol) and ethylamine (0.31 gm , 6.9 mmol) to yield **108** as rose colored powder (1.02 gm, 99%). mp: 182 °C from THF; ν_{max} (KBr disc) 3433, 3297, 1626, 1574, 1529 cm^{-1}; ^1H NMR (400 MHz, DMSO-d_6 , rotameric (Leftheris *et al*, 2004)): δ 1.05 (m, 3H, CH_3), 2.19 (s, 3H, CH_3), 3.20 (m, 2H, CH_2), 7.14-7.24 (m, 3H), 7.9 (br s,1H, NH), and 9.4 (br s,1H, NH) ppm. ^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 19.13 (CH_3), 22.29 (CH_3), 39.92 (CH_2), 129.5 (C), 130.27 (C), 134.48 (CH), 136.51 (CH), 142.46 (CH), 148.3 (C), 169.20 (C), 170.00 (C), and 172.95 (C). HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{12}H_{14}Cl_2N_5$: 298.06208 found: 298.06196.

6-Chloro-N-(5-chloro-2-methyl-phenyl)-N'-pentyl-[1,3,5]triazine-2,4-diamine (109)

prepared from **101** (1 gm, 3.5 mmol) and pentylamine (0.6 gm, 7 mmol) to yield **109** as rose powder (1.1 gm, 91%), mp: 155 C° from THF; ν_{max} (KBr disc) 3427, 3265, 1624, 1568, 1518 cm^{-1}; ^1H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 0.83 (m, 3H, CH_3), 1.22 (m, 4H, 2xCH_2), 1.45 (m, 2H, CH_2), 2.16 (m, 3H, CH_3), 3.13 (m, 2H, CH_2), 7.12-7.23 (m, 2H), 7.41 (m, 1H), 8.07 (br s,1H, NH), and 9.40 (br s,1H, NH) ppm. ^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 18.47 (CH_3), 22.27 (CH_3), 26.60 (CH_2), 33.23 (CH_2), 33.35 (CH_2), 39.99 (CH_2), 130.25 (C), 130.92 (C), 134.49 (CH), 136.47 (CH), 137.20 (C), 142.47 (CH), 169.19 (C), 170.16 (C), and 172.93 (C). HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{15}H_{20}Cl_2N_5$: 340.10903 found: 340.10908; Anal. Calcd for $C_{15}H_{19}Cl_2N_5$: C, 52.95; H, 5.63; N, 20.58 found: C 52.36; H, 5.48; N, 19.90.

6-Chloro-N-(5-chloro-2-methyl-phenyl)-N'-isopropyl-[1,3,5]triazine-2,4-diamine (110)

prepared from **101** (1.5 gm, 5.2 mmol) and isopropylamine (0.61 gm, 10.4 mmol). to yield **110** as rose colored powder (1.3 gm , 80%), mp: 220°C; ν_{max} (KBr disc) 3234, 2976, 1597, 1531 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6 , rotameric (Leftheris *et al*, 2004)): δ 1.09 (d, 6H, 2xCH_3), 2.16 (s, 3H, CH_3), 3.90 (m, 1H, CH), 7.09-7.22 (m, 2H), 7.41-7.60 (m, 1H), 7.78 (d,1H, NH), and 9.28 (br s,1H, NH) ppm. ^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 22.26 (CH_3), 26.98 (CH_3), 46.91 (CH), 129.57 (C), 130.98 (C), 134.51 (CH), 136.41 (CH), 137.2 (C), 142.48 (CH), 169.29 (C), 172.99 (C), and 173.60 (C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{13}H_{16}Cl_2N_5$: 312.07773 found: 312.07740.

6-Chloro-N-(4-chloro-phenyl)-N'-ethyl-[1,3,5]triazine-2,4-diamine(111) (Fahnenstich *et al*, 1974)

Prepared from **102** (1.0 gm, 3.63 mmol) and ethylamine (0.32 gm, 7.2 mmol) to yield **111** as white powder (0.97gm, 94%), mp: 220-222°C from THF; ν_{max} (KBr disc) 3259, 1637, 1579, 1527 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6 , rotameric (Leftheris *et al*, 2004)): δ 1.13 (m, 3H, CH_3), 3.25 (m, 2H, CH_2), 7.3-7.4 (m, 2H, 2xCH), 7.72-7.82 (m, 2H, 2xCH), 8.10 (br s,1H, NH), and 8.30 (br s,1H, NH) ppm. ^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 16.77 (CH_3), 37.82 (CH_2), 124.15 (2xCH), 128.80 (C), 131.1 (2xCH), 140.64 (C), 166.20 (C), 167.80 (C), and 170.10 (C). ppm; HRMS-ESI *m/z* $[M+Na]^+$ calcd for $C_{11}H_{11}C_{l2}N_5Na$: 306.02837 found: 306.02823; Anal. Calcd for $C_{11}H_{11}C_{l2}N_5$: C, 46.50; H, 3.90; N, 24.65 found: C 46.04; H 4.12; N, 23.28.

6-Chloro-N-(4-chloro-phenyl)-N'-pentyl-[1,3,5]triazine-2,4-diamine (112).

Prepared from **102** (1.0 gm, 3.63 mmol) and pentylamine (0.63 gm, 7.2 mmol) to yield **112** as white powder (1.1 gm, 94%). mp 190-192 °C from THF; ν_{max} (KBr disc) 3265, 3119, 1639, 1579 cm^{-1}; ^1H NMR (400 MHz, DMSO-d_6 , rotameric (Leftheris *et al*, 2004)): δ 0.82 (m, 3H, CH_3), 1.28 (m, 4H, 2xCH_2), 1.49 (m, 2H, CH_2), 3.23 (m, 2H, CH_2), 7.30 (m, 2H, 2xCH), 7.71 (m, 2H, 2xCH), 8.21 (br s,1H, NH), 10.14 (br s,1H, NH) ppm.. ^{13}C NMR (100 MHz, DMSO-d_6): δ 16.5 (CH_3), 24.8 (CH_2), 30.83 (CH_2), 31.12 (CH_2), 43.18 (CH_2), 124.11 (2xCH), 129.02 (C), 130.99 (2xCH), 140.66 (C), 166.29 (C), 167.99 (C), and 170.49 (C) ppm; HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{14}H_{18}Cl_2N_5$: 326.09338 found: 326.09312; Anal. Calcd for $C_{14}H_{17}Cl_2N_5$: C, 51.54; H, 5.25, N, 21.47 found: C, 51.16, H, 4.74, N, 20.57.

6-Chloro-N-(4-nitro-phenyl)-N'-pentyl-[1,3,5]triazine-2,4-diamine (113)

prepared from **103** (2 gm, 7.0 mmol) and pentylamine (1.2 gm, 13.7 mmol) to yield **113** as yellow powder (1.95 gm, 83 %). mp: 188-193 °C from THF; ν_{max} (KBr disc) 3300, 1635, 1541 cm^{-1}; ^1H NMR (400 MHz, DMSO-d_6 , rotameric (Leftheris *et al*, 2004)): δ 0.84 (m, 3H, CH_3), 1.29 (m, 4H, 2xCH_2), 1.54 (m, 2H, CH_2), 3.33 (m, 2H, CH_2), 7.93-8.10 (m, 2H), 8.13-8.17 (m, 2H), 8.39 (br s, 1H, NH), and 10.55 (br s, 1H, NH) ppm. ^{13}C NMR (100 MHz, DMSO-d_6): δ 16.56 (CH_3), 24.51 (CH_2), 30.75 (CH_2), 31.26 (CH_2), 43.32 (CH_2), 124.84 (2xCH), 127.28 (2xCH), 144.16 (C), 148.27 (C), 166.40 (C), 167.95 (C), and 170.61 (C) ppm; HRMS-ESI *m/z* $[M-H]^+$ calcd for $C_{14}H_{16}ClN_6O_2$: 335.10287 found: 335.10288.

1.3.2.3 Preparation of 4-[4-Alkylamino-6-(phenylamino)-[1,3,5]triazin-2-yl]-piperazine-1-carboxylic acid tert-butyl ester derivatives (114–115, scheme 1.3)

To a solution of the particular monochlorotriazine derivative (**104, 109**) in dioxane (50 ml) 1-*N*-BOC-piperazine (*ca.* 1.5 equivalents) and diisopropylethylamine (acid scavenger, 2 equivalents) were added successively. The reaction subsequently refluxed until completion as revealed by TLC. Distilled water (20 ml) was added to the reaction mixture and the resulting suspension was left at $4^{o}C$ over few days. The resulting suspensions were filtered and the solid residues were further purified by recrystallization from acetone/water (scheme 1.3).(Leftheris *et al*, 2004)

4-[4-Ethylamino-6-(2-methoxy-phenylamino)-[1,3,5]triazin-2-yl]-piperazine-1-carboxylic acid tert-butyl ester (114)

Prepared from **104** (0.7 gm, 2.5 mmol) and 1-*N*-BOC-piperazine (0.69 gm, 3.70 mmol) to yield **114** as white powder (0.75 gm, 70%), mp 161-162 from acetone/water; v_{max} (KBr disc) 3433, 3290, 2976, 1693 and 1575 cm^{-1}. ^{1}H NMR (400 MHz, DMSO-d_6): δ 1.08 (t, 3H, CH_3, J = 7.16), 1.42 (s, 9H, 3xCH_3), 3.26 (m, 2H, CH_2), 3.34 (br s, 4H, 2xCH_2), 3.68 (br s, 4H, 2xCH_2), 3.86 (s, 3H, OCH_3), 6.9-7.1 (m, 4H), 7.30-7.40 (br s, 1H, NH), and 8.28-8.29 (br s, 1H, NH) ppms; ^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 14.80 (CH_3), 28.03 (3xCH_3), 34.82 (CH_2), 42.44 (2x CH_2), 55.70 (OCH_3), 66.32 (2xCH_2), 78.99 (C), 110.52 (CH) 119.76 (CH), 120.34 (CH), 121.99 (CH), 128.61 (C), 148.14 (C), 153.87 (C=O), 163.76 (C), 164.49 (C), and 165.35 (C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{21}H_{32}N_7O_3$: 430.25611 found: 430.25493.

4-[4-(5-Chloro-2-methyl-phenylamino)-6-pentylamino-[1,3,5]triazin-2-yl]-piperazine-1-carboxylic acid tert-butyl ester (115)

Prepared from **109** (0.5 gm, 1.47 mmol) and 1-*N*-BOC-piperazine (0.27 gm, 1.47 mmol) to yield **115** as brown powder (0.61 gm, 85%), mp: 123-124 from acetone/water; ν_{max} (KBr disc) 3447, 3364, 1682, 1606, 1545 cm^{-1}. ^{1}H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): 0.84 (m, 3H, CH_3), 1.26 (m, 4H, $2xCH_2$), 1.42 (s, 9H, $3xCH_3$), 1.46 (m, 2H, CH_2), 2.18 (s, 3H, CH_3, splitted due to rotamers), 3.17 (m, 2H, CH_2), 3.21 (m, 4H, $2xCH_2$), 3.65 (m, 4H, $2xCH_2$), 6.89-7.18 (m, 3H, 3xCH), 7.62 (br s, 1H, NH), and 8.15 (br s, 1H, NH) ppm;^{13}C NMR (100 MHz, DMSO-d_6 rotameric (Leftheris *et al*, 2004)): 13.91 (CH_3), 17.51 (CH_3), 21.85 (CH_2), 28.01 ($3xCH_3$), 28.61 ($2xCH_2$), 28.75 (CH_2), 29.11 (CH_2), 39.85 (CH_2), 42.32 ($2xCH_2$), 78.99 (C), 122.78 (CH), 123.73 (CH), 129.21 (C), 129.56 (C), 131.26 (CH), 139.12 (C), 153.87 (C=O), 164.38 (C), 164.50 (C), and 165.66 (C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{24}H_{37}ClN_7O_2$: 490.26918 found: 490.26949.

1.3.2.4 Preparation of 4-[4-alkylamino-6-(phenylamino)-[1,3,5]triazin-2-yl]-piperazin-1-ium; chloride derivatives (116–117, scheme 1.3)

The particular *N*-protected-piperazinyl-triazine derivative (**114**, **115**) was dissolved in methanol (25 ml) and magentically stirred at 0 °C for 30 min. Subsequently, aqueous HCl solution (6 ml, 6 N) was added dropwise to the methanolic solution at 0^{o}C over 70 min. Thereafter, the resulting mixture was warmed to ambient temperature over 3h, then was slowly heated to an internal temperature of 50 °C and left on a stirrer for 24 hr. The resulting aqueous solution was cooled to 0°C and made alkaline (pH = 14) by addition NaOH solution (10% w/v). The resulting alkaline solution was extracted with chloroform (3x25ml) and the organic phases were combined and dried over

magnesium sulfate. The solvent is left to evaporate *in vacuo* to afford the yellow oil. Finally, the oil was dissolved in acetone (10 ml), filtered and, then precipitated by acetonic HCl to yield the final product as white powder.

N-Ethyl-N'-(2-methoxy-phenyl)-6-piperazin-1-yl-[1,3,5]triazine-2,4-diamine (116)

Prepared from compound **114** (0.5 gm, 1.17 mmol) to yield **116** as white powder (0.40 gm, 78%), mp: 256 (Decomp.); v_{max} (KBr disc) 3431, 1639, 1595 cm^{-1}; ^{1}H NMR (300 MHz, DMSO-d_6, as rotameric (Leftheris *et al*, 2004)): ^{1}H NMR (300 MHz, DMSO-d_6): □ 1.10 (t, 3H, CH_3, J = 7.0), 3.13 (m, 4H, 2xCH_2), 3.31 (m, 2H, CH_2), 3.79 (s, 3H, OCH_3), 3.96 (m, 4H, 2xCH_2), 6.92-7.19 (m, 4H), 7.74 (br s, H, NH), 8.76 (br s, H, NH), and 9.79 (br s, 1H, NH) ppms. HRMS-ESI m/z $[M+H]^+$ calcd for $C_{16}H_{24}N_7O$: 330.20368, found 330.20247; Anal. calcd for $C_{16}H_{23}N_7O$.3HCl.: C, 43.80; H, 5.97; N, 22.35. Found: C, 42.45; H, 5.61; N, 21.57.

4-[4-(5-Chloro-2-methyl-phenylamino)-6-pentylamino-[1,3,5]triazin-2-yl]-piperazine.HCl (117)

Prepared from **115** (0.5 gm, 1.02 mmol) to yield **117** as white powder, mp 124 ^{o}C from acetone (Decomp.); v_{max} (KBr disc) 3421, 1602, 1572, 1548 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): □ 0.84 (m, 3H. CH_3), 1.23 (m, 4H, 2xCH_2), 1.46 (m, 2H, CH_2), 2.18 (s, 3H, CH_3, splitted due to rotamers), 3.18 (m, 2H, CH_2), 3.69 (m, 4H, 2xCH_2), 6.87-7.18 (m, 3H), 7.63 (br s, 1H, NH), and 8.11 (br s, 1H, NH) ppm; ^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): □ 13.92 (CH_3), 17.52 (CH_3), 21.84 (CH_2), 28.60 (CH_2), 29.11 (CH_2), 39.72 (CH_2), 42.34 (2xCH_2), 44.43 (2xCH_2), 122.73 (CH), 123.67 (CH), 129.54 (C), 130.10 (C), 131.29 (CH) 139.25 (C), 164.37 (2xC), and 165.7 (C) ppm. HRMS-ESI m/z $[M+H]^+$ calcd for

$C_{19}H_{29}ClN_7$: 390.21675 found: 390.21653. Anal. calcd for $C_{19}H_{28}ClN_7.H_2O.3HCl$.: C, 44.11; H, 6.43; N, 18.95. Found: C, 43.37; H, 6.20; N, 17.85.

1.3.2.5 Preparation of 6-[4-(2-Dimethylamino-ethyl)-piperazin-1-yl]-N-(phenyl)-N'-alkyl-[1,3,5]triazine-2,4-diamine (118–127, scheme 1.4)

To a magnetically stirred solution of the particular mono-chloro-diamino-triazine derivative (**104-113**) in THF (20 ml), 1-[2-(dimethylamino)ethyl]piperazine (1.5-3 equivalents) and triethylamine or diisopropylethylamine (acid scavenger, 2 equivalent) were added sequentially. The resulting mixture was stirred at room temperature until completion as revealed by TLC (roughly over two days). Subsequently, the reaction mixture was filtred and the filtrate was dried *in vacuo*. The residue was dissolved in acetone, filtered and precipitated by acetonic HCl to yield powdered product (scheme 1.4) (Leftheris *et al*, 2004).

6-[4-(2-Dimethylamino-ethyl)-piperazin-1-yl]-N-ethyl-N'-(2-methoxy-phenyl)-[1,3,5]triazine-2,4-diamine (118)

Prepared from **104** (0.40 gm, 1.43 mmol) and 1-[2-(dimethyl amino) ethyl] piperazine (0.36 gm, 2.29 mmol) and triethylamine (0.29 gm, 2.87mmol) to yield **118** as white powder (0.4 gm, 48%), mp: 230°C (Decomp.) from acetone; ν_{max} (KBr disc) 3421, 1620, 1593, 1547 cm^{-1}; ^1H NMR (400 MHz, DMSO-d6, rotameric (Leftheris *et al*, 2004), protonated sample): δ 1.14 (m, 3H, CH_3), 2.83 (s, 6H, $N(CH_3)_2$), 3.04 (m, 2H, CH_2), 3.19 (m, 4H, 2xCH_2), 3.62 (s, 3H, OCH_3), 3.83 (m, 8H, 4xCH_2), 6.99-7.89 (m, 4H), 8.92 (br s, 1H, NH), and 10.07 (br s, 1H, NH) ppm;^{13}C NMR (100 MHz, DMSO-*d6*): δ 14.03 (CH_3), 35.37 (CH_2), 42.33 ($N(CH_3)_2$), 49.33 (CH_2), 50.29 (CH_2), 52.75 (2xCH_2), 55.81 (OCH_3), 57.90 (2xCH_2), 111.70 (br s, 2xCH), 120.4 (br s, 2xCH), 122.0 (br s, C) 126.0 (br s, C), 152.0 (br s, C), 156 (br s, C), and 161.0 (br s,

C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{20}H_{33}N_8O$: 401.27718 found: 401.27718. Anal. Calcd for $C_{20}H_{32}N_8O.5HCl$: C, 41.22; H, 6.40, N, 19.23 found: C, 41.74, H, 6.22, N, 19.03.

6-[4-(2-Dimethylamino-ethyl)-piperazin-1-yl]-N-isopropyl-N'-(2-methoxy-phenyl)-[1,3,5]triazine-2,4-diamine (119)

Prepared from **105** (0.50 gm, 1.7 mmol) and 1-[2-(dimethylamino)ethyl] piperazine (0.35 gm, 2.23 mmol) and triethylamine (0.29 gm, 2.9 mmol) to yield **119** as brown powder (0.62 gm, 56%), mp: 240 (Decomp.) from acetone, ^{1}H NMR (400 MHz, DMSO-*d6* and TFA-*d*, rotameric (Leftheris *et al*, 2004), protonated sample): 1.1 (d, 6H, 2xCH_3), 2.86 (s, 6H, $N(CH_3)_2$), 3.49-3.61 (m, 12H, 6xCH_2), 3.83 (s, 3H, OCH_3), 4.14 (m, 1H, CH), 6.98-7.89 (m, 4H), 9.10 (br s, 1H, NH, *d*-Exchangeable), 10.0 (br s, 1H, NH, *d*-Exchangeable) ppm;^{13}C NMR (100 MHz, DMSO-d6, rotameric (Leftheris *et al*, 2004)): 21.45 (2xCH_3), 39.82 (CH_2), 42.30 ($N(CH_3)_2$), 42.80 (CH), 48.26 (CH_2), 49.60 (2xCH_2), 49.68 (2xCH_2), 55.59 (OCH_3), 111.38 (CH), 111.84 (CH), 113.45(C), 116.31(C), 120.42 (CH), 125.70 (CH), 157.63 (C), 158.02 (C), and 158.40 (C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{21}H_{35}N_8O$: 415.29283 found: 415.29192. Anal. Calcd for $C_{21}H_{34}N_8O.5H_2O.4HCl$: C, 38.78; H, 7.44, N, 17.23 found: C, 37.86, H, 7.03, N, 16.51.

N-(5-chloro-2-methyl-phenyl)-6-[4-(2-dimethylamino-ethyl)-piperazin-1-yl]-N'-ethyl-[1,3,5]triazine-2,4-diamine (120)

Prepared from **108** (0.40 gm, 1.34 mmol) and 1-[2-(dimethylamino)ethyl] piperazine (0.317 gm, 2.02 mmol) and triethylamine (0.27 gm, 2.6 mmol) to yield **120** as white powder (0.278 gm, 39%), mp 240°C (Decomp.) from acetone; ν_{max} (KBr disc) 3649, 3800, 1655, 1624 cm^{-1}; ^{1}H NMR (300 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004), protonated sample): 1.07 (t, 3H, CH_3, *J*=7), 2.21 (s, 3H, CH_3), 2.79 (s, 6H,

$N(CH_3)_2$), 3.34 (m, 8H, 4xCH_2), 3.58 (m, 6H, 3xCH_2), 7.24-7.73 (m, 3H), 8.80 (br s, 1H, NH), and 10.30 (br s, 1H, NH) ppm; ^{13}C NMR (75 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 14.57 (CH_3), 18.06 (CH_3), 35.87 (CH_2), 42.75 (2xCH_3, $N(CH_3)_2$), 49.79 (2xCH_2), 50.25 (2xCH_2), 50.77 (br s, 2x CH_2), 124.56 (CH), 125.54 (CH), 130.58 (C), 131.90 (C), 132.44 (CH), 136.07 (C), 155.65 (C), 160.0 (C) and 161.55 (C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{20}H_{32}ClN_8$: 419.24330 found: 419.24301. Anal. calcd for $C_{20}H_{31}ClN_8$.3HCl: C, 45.46; H, 6.49; N, 21.21, Found: C, 44.92; H, 6.49; N, 20.30.

N-(5-chloro-2-methyl-phenyl)-6-[4-(2-dimethylamino-ethyl)-piperazin-1-yl]-N'-pentyl-[1,3,5]triazine-2,4-diamine (121)

Prepared from **109** (0.40 gm, 1.18 mmol) and 1-[2-(dimethylamino)ethyl] piperazine (0.45 gm, 2.9 mmol) and triethylamine (0.24 gm, 2.37 mmol) to yield **121** as white powder (0.067 gm, 9%), mp 71-72°C from acetone (unprotonated sample); ν_{max} (KBr disc) 3431, 1577, 1543, 1500 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004), unprotonated sample): δ 0.84 (m, 3H, CH_3), 1.23 (m, 4H, 2xCH_2), 1.46 (m, 2H, CH_2), 2.13 (s, 6H, $N(CH_3)_2$), 2.34 (s, 3H, CH_3), 2.38 (m, 8H, 4xCH_2), 3.19 (m, 2H, CH_2), 3.64 (m, 4H, 2xCH_2), 6.83-7.18 (m, 3H) 7.65 (br s, 1H, NH), and 8.07 (br s, 1H, NH) ppm;^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 13.92 (CH_3), 17.51 (CH_3), 21.83 (CH_2), 28.61 (CH_2), 29.14 (CH_2), 39.53 (CH_2), 42.55 (CH_2), 45.51 (2xCH_3, $N(CH_3)_2$), 52.93 (2xCH_2) 55.90 (CH_2), 56.58 (2xCH_2), 122.66 (CH), 123.59 (CH), 129.06 (C), 129.54 (C), 131.28 (CH), 139.27 (C), 164.36 (C), 1 64.50 (C), and 165.77 (C) ppm. HRMS-ESI *m/z* $[M+H]^+$ calcd for $C_{23}H_{38}ClN_8$: 461.29025 found: 461.29033. Anal. calcd for $C_{23}H_{37}ClN_8$.5HCl: C, 42.94 Found: C, 41.49.

N-(5-chloro-2-methyl-phenyl)-6-[4-(2-dimethylamino-ethyl)-piperazin-1-yl]-N'-isopropyl-[1,3,5]triazine-2,4-diamine (122)

Prepared from **110** (0.50 gm, 1.6 mmol) and 1-[2-(dimethylamino)ethyl] piperazine (0.38 gm, 2.42 mmol) and triethylamine (0.32 gm, 3.17 mmol) to yield **122** as white powder (0.44 gm, 46%), mp: 245°C (Decomp.) from acetone; ν_{max} (KBr disc) 3447, 1655, 1626, 1593 cm^{-1}, ^{1}H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004), protonated sample): δ 1.18 (d, 6H, 2xCH_3, J=8), 2.24 (s, 3H, CH_3), 2.82 (s, 6H, $N(CH_3)_2$), 3.43 (m, 12H, 6xCH_2), 4.14 (m, 1H, CH), 7.24-7.28 (m, 3H), 7.59 (br s, H, NH), and 8.10 (br, s, H, NH) ppm. HRMS-ESI m/z $[M+H]^+$ calcd for $C_{21}H_{34}ClN_8$: 433.25895 found: 433.25228; Anal. Calcd for $C_{21}H_{33}ClN_8.4HCl.H_2O$: C, 42.26; H, 6.59; N, 18.77 found: C, 42.15; H, 6.41; N 18.33.

6-[4-(2-Dimethylamino-ethyl)-piperazin-1-yl]-N-isopropyl-N'-(2-methoxy-4-methyl-phenyl)-[1,3,5]triazine-2,4-diamine (123)

Prepared from **107** (0.4 gm, 1.30 mmol) and 1-[2-(dimethylamino)ethyl] piperazine (0.53 gm, 3.37 mmol) and triethylamine (0.23 gm, 2.27 mmol) to yield **123** as white powder (0.22 gm, 28%), mp: 214°C (Decomp.) from acetone; ν_{max} (KBr disc) 3435, 1629, 1599, 1535 cm^{-1}, ^{1}H NMR (300 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004), protonated sample): δ 1.20 (d, 6H, 2xCH_3), 2.21 (s, 3H, CH_3), 2.79 (s, 6H, $N(CH_3)_2$), 3.56-3.75 (m, 12H, 6xCH_2), 3.71(s, 3H, OMe), 4.10 (m, 1H), 6.92 (m, 3H), 7.48 (br s,1H, NH), and 8.70 (br s,1H, NH) ppm. ^{13}C NMR (75 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 20.82 (CH_3) 22.34 (2xCH_3), 41.05 (CH_2), 42.81 (2xCH_3, $N(CH_3)_2$), 43.33 (CH), 49.94 (CH_2), 50.41 (2xCH_2) 50.91 (2xCH_2), 56.36 (OCH_3), 111.83 (CH), 123.85 (CH), 125.94 (CH), 127.50 (C), 129.50 (C), 148.79 (C), 150.0 (C), 155.46 (C), and 162.50 (C) ppm. HRMS-ESI m/z $[M+H]^+$ calcd for

$C_{22}H_{37}N_8O$: 429.30848 found: 429.30807. Anal. Calcd for $C_{22}H_{37}ClN_8O.5HCl$: C, 43.25; H, 6.76, N, 18.34 found: C 44.14, H 6.82, N, 18.37.

6-[4-(2-Dimethylamino-ethyl)-piperazin-1-yl]-N-(2-methoxy-4-methyl-phenyl)-N'-pentyl-[1,3,5]triazine-2,4-diamine (124)

Prepared from **106** (0.5 gm, 1.5 mmol) and 1-[2-(dimethylamino)ethyl] piperazine (0.69 gm, 4.4 mmol) and diisopropylethylamine (0.34 gm, 2.6 mmol) to yield **124** as cream colored powder (0.44 gm, 64%). mp: 240 (Decomp.) from acetone; v_{max} (KBr disc) 3425, 1616, 1506 cm^{-1}, ^1H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004), unprotonated sample): δ 0.85 (m, 3H, CH_3), 1.29 (m, 4H, 2xCH_2), 1.51 (m, 2H, CH_2), 2.24 (s, 6H, $N(CH_3)_2$), 2.25 (s, 3H, CH_3), 2.36 (m, 8H, 4xCH_2), 3.24 (m, 2H, CH_2), 3.60 (m, 4H, 2xCH_2), 3.81 (s, 3H, OCH_3), 6.73-7.03 (m, 3H), 7.30 (br s, 1H, NH) and 8.21 (br s, 1H, NH) ppm; ^{13}C NMR (100 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004)): δ 13.90, (CH_3), 20.73 (CH_3), 21.84 (CH_2), 28.64 (CH_2), 29.01 (CH_2), 39.52 (CH_2), 42.64 (CH_2), 45.49 ($N(CH_3)_2$), 52.91 (2xCH_2), 55.73 (OCH_3), 55.90 (CH_2), 56.57 (2xCH_2), 110.23 (CH), 120.29 (CH), 121.79 (CH), 128.40 (C) 128.77 (C), 145.80 (C), 163.70 (C), 164.32 (C) and 165.56 (C) ppm. HRMS-ESI *m/z* $[M+Na]^+$ calcd for $C_{24}H_{40}N_8NaO$: 479.32173 found: 479.32172.

N-(4-Chloro-phenyl)-6-[4-(2-dimethylamino-ethyl)-piperazin-1-yl]-N'-ethyl-[1,3,5]triazine-2,4-diamine (125)

Prepared from **111** (0.4 gm, 1.41 mmol) and 1-[2-(dimethylamino)ethyl] piperazine (0.66 gm, 4.2 mmol) and triethylamine (0.28 gm, 2.8 mmol) to yield **125** as white powder (0.3 gm, 52%). mp: 232 (Decomp.) from acetone. v_{max} (KBr disc) 3437, 1624, 1591 cm^{-1}, ^1H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004), unprotonated sample): δ 1.05 (t, 3H, CH_3, J=8), 2.10 (s, 6H, $N(CH_3)_2$), 2.35 (m, 8H,

4xCH_2), 3.25 (m, 2H, CH_2), 3.60 (m, 4H, 2xCH_2), 6.84 (br s, 1H, NH), 7.28 (m, 2H) 7.71 (m, 2H), and 8.94 (br s, 1H, NH) ppm;^{13}C NMR (100 MHz, DMSO-d6, rotameric (Leftheris *et al*, 2004)): □ 17.50 (CH_3), 37.45 (CH_2), 45.28 (2xCH_3, $N(CH_3)_2$), 48.07 (2xCH_2), 55.55 (2xCH_2), 58.39 (CH_2), 59.02(CH_2), 123.46 (2xCH), 127.50 (C), 130.76 (2xCH), 142.20, 166.20 (C), 167.00 (C), and 167.80 (C) ppm. HRMS-ESI *m/z* [*M*+H]$^+$ calcd for $C_{19}H_{30}ClN_8$: 405.22765 found: 405.22778. Anal. Calcd for $C_{19}H_{29}ClN_8$. 5HCl : C, 38.86; H, 5.84; N, 19.08 found: C 39.67, H 5.58, N, 19.14.

N-(4-Chloro-phenyl)-6-[4-(2-dimethylamino-ethyl)-piperazin-1-yl]-N'-pentyl-[1,3,5]triazine-2,4-diamine (126)

Prepared from **112** (0.4 gm, 1.2 mmol) 1-[2-(dimethylamino)ethyl] piperazine (0.56 gm, 3.6 mmol) and diisopropylethylamine (0.05 gm, 0.38 mmol) to yield **126** as off-white powder (0.24 gm, 31%). mp: 224°C (Decomp.) from acetone., ^{1}H NMR (400 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004), protonated sample): □ 0.87 (m, 3H. CH_3), 1.30 (m, 4H, 2xCH_2), 1.53 (m, 2H, CH_2), 2.83 (s, 6H, $N(CH_3)_2$), 3.33 (m, 2H,CH_2), 3.60-3.79 (m, 12, 6xCH_2), 7.39 (m, 2H), 7.64 (m, 2H), 8.30 (br s, 1H, NH) and 11.0 (br s, 1H, NH) ppm; ^{13}C NMR (100 MHz, DMSO-d_6): □ 13.89 (CH_3), 21.75 (CH_2), 28.20 (CH_2), 28.39 (CH_2), 40.47 (CH_2), 42.38 ($N(CH_3)_2$), 49.51 (2xCH_2), 49.9 (2xCH_2), 50.50 (CH_2), 50.56 (CH_2), 122.02 (2xCH), 124.11(C), 128.68 (2xC), 140.60 (C), 155.34 (C), 156.82(C), and 164.04 (C) ppm. HRMS-ESI *m/z* [*M*+H]$^+$ calcd for $C_{22}H_{36}ClN_8$: 447.27460 found: 447.27430. Anal. Calcd for $C_{22}H_{35}ClN_8$.5HCl : C, 41.99; H, 6.41; N, 17.81 found: C 41.82, H 6.37, N, 17.32.

6-[4-(2-Dimethylamino-ethyl)-piperazin-1-yl]-N-(4-nitro-phenyl)-N'-pentyl-[1,3,5]triazine-2,4-diamine (127)

Prepared from **113** (0.4 gm, 1.19 mmol) and 1-[2-(dimethylamino)ethyl] piperazine (0.56 gm, 3.56 mmol) and diisopropylethylamine (0.05 gm, 0.38 mmol) to yield **127** as off-white powder (0.4 gm, 51%). mp: 237°C (Decomp.) from acetone; ν_{max} (KBr disc) 3464, 1662, 1626, 1585 cm^{-1}; ^{1}H NMR (300 MHz, DMSO-d_6, rotameric (Leftheris *et al*, 2004), protonated sample): δ 0.84 (m, 3H. CH_3), 1.28 (m, 4H, 2xCH_2), 1.51 (m, 2H, CH_2), 2.81 (s, 6H, $N(CH_3)_2$), 3.57 (m, 4H, 2xCH_2), 3.96 (m, 10H, 5xCH_2), 7.93 (d, 2H, *J*=9.0), 7.68 (br s, 1H, NH), 8.12 (d, 2H, *J*=9.0), and 10.22 (br s, 1H, NH) ppm; ^{13}C NMR (75 MHz, DMSO-*d6*): δ 14.40 (CH_3), 22.32 (CH_2), 29.07 (2xCH_2), 40.77 (CH_2), 42.93 ($N(CH_3)_2$), 50.15 (2xCH_2), 50.61 (2xCH_2), 51.24 (2xCH_2), 119.63 (2xCH), 125.21 (2xCH), 141.64 (C), 146.81 (C), 163.21 (C), 163.78 (C) and 164.20 (C) ppm. HRMS-ESI m/z [M+H]+ calcd for $C_{22}H_{36}N_9O_2$: 458.29865 found: 458.29756. Calcd for $C_{22}H_{35}N_9O_2.3H_2O.4HCl$: C, 40.19; H, 6.90; N, 19.17 found: C 40.14, H 6.73, N, 18.11.

1.3.3 *In vitro* experimental studies

1.3.3.1 Materials of Bioassay

All of the chemicals used in these experiments were of reagent grade and obtained from commercial suppliers. NCI samples were kindly provided by the National Cancer Institute. CaMKII drug discovery kit was purchased from Cyclex (Japan), the standard inhibitor KN-62 (table 1.11), water and DMSO for bioanalysis were all purchased form Sigma-Aldrich (USA).

1.3.3.2 Preparation of hit compounds for *In vitro* assay

The tested compounds were provided as dry powders in variable quantities (5 mg to 10 mg). They were initially dissolved in DMSO to give stock solutions of fixed concentrations. Subsequently, they were diluted to the required concentrations with deionized water for bioassay

1.3.3.3 Quantification of CaMKII activity in a spectrophotometric assay

The activity of the in silico hits were quantified by CaMKII drug discovery kit (Cyclex, Japan). The 96-well plate of the kit is pre-coated with “Syntide-2” phosphorylated by CaMKII. The detector antibody (MS-6E6) is conjugated to horseradish peroxidase and specifically detects the phosphorylated form of “Syntide-2”.

To perform the assay, the tested hits were first diluted with the kinase buffer provided with the kit. The hits' solutions were then pipetted into assay wells to yield final concentrations of 0.1, 1.0 and 10 μM and occasionally at 0.2 and 0.02 μM. Subsequently, CaMKII was added to each well as aqueous solution (0.01 units in 10 μL) and allowed to phosphorylate the bound substrate in the presence of $CaCl_2$, Calmodulin and ATP. The amount of phosphorylated substrate was measured by allowing it to bind with the detector antibody conjugated to horseradish peroxidase, which then catalyzes the conversion of the chromogenic substrate tetra-methylbenzidine (TMB) from a colorless solution to a blue solution. The color is quantified by spectrophotometry and reflects the relative activity of CaMKII. Samples' absorbances were determined at a wavelength of 650 nm using a plate reader

(Bio-Tek instruments ELx 800, USA). Inhibition of CaMKII was calculated as percent activity of the uninhibited CaMKII enzyme control. KN-62 was used as positive control. Negative controls were prepared by adding the enzyme to the kinase reaction buffer only (Cyclex CaM kinase II Assay Kit (Cat# CY-1173) Users' Manual, 2009).

1.4. Results and Discussion

CATALYST models drug-receptor interaction using information derived only from the ligand structure. It enables automatic pharmacophore construction by using a collection of molecules with activities ranging over a number of orders of magnitude. HYPOGEN identifies a 3D array of a maximum of five chemical features common to active training molecules, which provides a relative alignment for each input molecule consistent with their binding to a proposed common receptor site. The chemical features considered can be hydrogen bond donors and acceptors (HBDs and HBAs), aliphatic and aromatic hydrophobes (Hbic), positive and negative ionizable (PosIon and NegIon) groups and aromatic planes (RingArom). The conformational flexibility of training ligands is modeled by creating multiple conformers, judiciously prepared to emphasize representative coverage over a specified energy range. CATALYST pharmacophores have been used as 3D search queries for screening NCI database (Taha *et al.*, 2008, a,b&c; Taha *et al.*, 2007; Taha *et al.*, 2010; Al-masri *et al*; 2008; Al-Nadaf *et al*; 2010; Abu-Hammad *et al*, 2009; Abu Khalaf *et al*, 2010; Al-Sha'er *et al*, 2010, a&b; Abdula *et al*, 2011; Abu Khalaf *et al*, 2011).

Different hypotheses were generated for a series of CaMKII inhibitors. A total of 88 compounds were used in this study (Table 1.1) (Levy *et al.,* 2008, a&b; Lu *et al.,* 2008; Mavunkel *et al.,* 2008). Nine training subsets were selected from the collection

(Table 1.2). Each subset consisted of inhibitors of wide structural diversity. The biological activity in the training subsets spanned from 3.5 to 4.0 orders of magnitude.

1.4.1. Data Mining and Conformational Coverage

The literature was surveyed to collect many reported structurally diverse CaMKII inhibitors (**1-88**, Table 1.1) (Levy *et al.,* 2008, a&b; Lu *et al.,* 2008; Mavunkel *et al.,* 2008). The 2D structures of the inhibitors were imported into CATALYST and converted automatically into 3D single conformer representations. The structures were used as starting points for conformational analyses and in the determination of various molecular descriptors for QSAR modeling.

The conformational space of each inhibitor was extensively sampled utilizing the poling algorithm employed within the CONFIRM module of CATALYST (Sutter *et al.,* 2000). Conformational coverage was performed employing the "Best" module to ensure extensive sampling of conformational space. Pharmacophore generation and pharmacophore-based search procedures are known for their sensitivity to inadequate conformational sampling within the training compounds prompting our extensive conformational sampling (Sheridan *et al*., 2002).

1.4.2 Exploration of CaMKIIδ Pharmacophoric Space

The fact that we have an informative list of 88 CaMKIIδ inhibitors of evenly spread bioactivities over 3.5 orders of magnitude, prompted us to employ CATALYST-HYPOGEN to identify possible pharmacophoric binding modes assumed by CaMKδII inhibitors (Levy *et al.,* 2008, a&b; Lu *et al.,* 2008; Mavunkel *et al.,* 2008).

HYPOGEN implements an optimization algorithm that evaluates large number of potential binding models for a particular target through fine perturbations to hypotheses that survived the constructive and subtractive phases of the modeling algorithm (see section **1.3.1.2.3 Pharmacophoric Hypotheses Generation** in **Experimental**) (Li *et al.*, 2000). The extent of the evaluated pharmacophoric space is reflected by the configuration (Config.) cost calculated for each modeling run. It is generally recommended that the Config. cost of any HYPOGEN run not to exceed 17 (corresponding to 2^{17} hypotheses to be assessed by CATALYST) to guarantee thorough analysis of all models (Sutter *et al.,* 2000). The size of the investigated pharmacophoric space is a function of training compounds, selected input chemical features and other CATALYST control parameters (Sutter *et al.,* 2000).

Restricting the extent of explored pharmacophoric space should improve the efficiency of optimization via allowing effective evaluation of limited number of pharmacophoric models. On the other hand, rigorous restrictions imposed on the pharmacophoric space might reduce the possibility of discovering optimal pharmacophoric hypotheses, as they might occur outside the “boundaries” of the pharmacophoric space.

Therefore, we decided to explore the pharmacophoric space of CaMKIIδ inhibitors under reasonably imposed "boundaries" through 68 HYPOGEN automatic runs and employing nine carefully selected training subsets: subsets **I-IX** in table 1.2. The training compounds in these subsets were selected in such away to guarantee maximal 3D diversity and continuous bioactivity spread over more than 3.5 logarithmic cycles. We gave special emphasis to the 3D diversity of the most active compounds in each training subset (Table 1.2) because of their significant influence on the extent of the evaluated pharmacophoric space during the constructive phase of HYPOGEN

algorithm (see section **1.3.1.2.3 Pharmacophoric Hypotheses Generation** in **Experimental**) (Li *et al*, 2000).

Guided by our rationally restricted pharmacophoric exploration concept, we restricted the software to explore pharmacophoric models incorporating from zero to one PosIon or NegIon features, from zero to three HBA, Hbic, or RingArom features instead of the default range of zero to five, as shown in Table 1.3. Furthermore, we instructed HYPOGEN to explore only 4- and 5-featured pharmacophores, i.e., ignore models of lesser number of features in order to further narrow the investigated pharmacophoric space and to represent the feature-rich nature of known CaMKIIδ ligands (as shown in table 1.3).

In each run, the resulting binding hypotheses were automatically ranked according to their corresponding "total cost" value, which is defined as the sum of error cost, weight cost and configuration cost (see section see section **1.3.1.2.4 Assessment of the Generated Hypotheses** in **Experimental**) (Bersuker *et al*. 2000; CATALYST 4.11 users' manual, 2005; Kurogi and Güner, 2001; Li *et al*., 2000; Poptodorov *et al*. 2006; Sutter *et al*. 2000). Error cost provides the highest contribution to total cost and it is directly related to the capacity of the particular pharmacophore as 3D-QSAR model, i.e., in correlating the molecular structures to the corresponding biological responses (Bersuker *et al*. 2000; CATALYST 4.11 users' manual, 2005; Kurogi and Güner, 2001; Li *et al*., 2000; Poptodorov *et al*. 2006; Sutter *et al*. 2000). HYPOGEN also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are normally distributed about their mean. Accordingly, the greater the difference from the null hypothesis cost (residual cost, Table 1.4) the more likely that the hypothesis does not reflect a chance correlation. An additional validation technique based on Fischer's randomization test

(Fisher, 1966) was recently introduced into CATALYST: Cat. Scramble (CATALYST 4.11 users' manual, 2005). In this test the biological data and the corresponding structures are scrambled several times and the software is challenged to generate pharmacophoric models from the randomized data. The confidence in the parent hypotheses (i.e., generated from unscrambled data) is lowered proportional to the number of times the software succeeds in generating binding hypotheses from scrambled data of apparently better cost criteria than the parent hypotheses (see section 4.1.5 Assessment of the Generated Hypotheses in Experimental) (Bersuker *et al*. 2000; CATALYST 4.11 users' manual, 2005; Kurogi and Güner, 2001; Li *et al*., 2000; Poptodorov *et al*. 2006; Sutter *et al*. 2000).

Eventually, 680 pharmacophore models emerged from 68 automatic HYPOGEN runs, out of which only 677 models illustrated Cat. scramble confidence levels ≥85%. These successful models were clustered and the best representatives (134 models, Table 1.4, see section 1.3.1.2.5 Clustering of the Generated Pharmacophore Hypotheses) were used in subsequent QSAR modeling. Clearly from table 1.4, representative models shared comparable features and acceptable statistical success criteria. Emergence of several statistically comparable pharmacophore models suggests the ability of CaMKIIδ ligands to assume multiple pharmacophoric binding modes within the binding pocket. Therefore, it is quite challenging to select any particular pharmacophore hypothesis as a sole representative of the binding process.

Table 1.4: Success criteria of representative pharmacophoric hypotheses (cluster centers).

RUN[a]	Hypotheses[b]	Features	Cost				R[d]	Global R[e]	F-stat[f]	Cat. Scramble
			Config.	Total	Null	Residual[c]				
1	5	HBA, Hbic, PosIon, RingArom	18.0	108.8	160.1	50.5	0.93	0.16	16.39	95%
	10	Hbic, PosIon, 2xRingArom	18.0	109.6	160.1	51.3	0.92	0.17	17.7	95%
2	8	HBD, 3xHbic.	17.0	107.2	160.1	52.9	0.95	0.04	3.32	95%
4	6	HBD, 3xHbic, RingArom.	13.8	107.7	160.1	52.3	0.91	0.1	9.32	95%
	9	2xHBD, 3xHbic.	13.8	108.8	160.1	51.2	0.91	0.12	11.45	95%
5	1	3xHbic, PosIon, , RingArom, 5 Exvol	18.0	102.8	160.1	57.3	0.97	0.15	15.53	95%
	4	HBA , 3xHbic, PosIon, 3xExvol	18.0	108.1	160.1	52	0.93	0.07	6.86	95%
	6	HBD, 2xHbic, PosIon.	18.0	108.9	160.1	51.2	0.92	0.23	25.89	95%
	9	HBA, HBD, 2xHbic, PosIon, 3x Exvol	18.0	111.6	160.1	48.5	0.91	0.12	11.19	95%
7	1	HBD, 2xHbic, PosIon,RingArom,7xExvol	15.4	99.6	160.1	60.4	0.97	0.26	29.97	95%
	4	2xHBD, 2xHbic, PosIon, 5x Exvol	15.4	102.8	160.1	57.3	0.95	0.22	23.92	95%
	10	HBA, HBD, 2xHbic, PosIon, 2x Exvol	15.4	105.7	160.1	54.4	0.93	0.23	26.34	95%
8	2	2xHBD, 2xHbic, PosIon, 5x Exvol	13.8	101.4	160.1	58.7	0.95	0.02	2.1	95%
	3	3xHbic, PosIon, RingArom, 1x Exvol	13.8	105	160.1	55.1	0.93	0.2	22.15	95%
	4	HBA, 2xHbic,PosIon,RingArom, 5xExvol	13.8	105.5	160.1	54.6	0.93	0.18	19.07	95%
	8	HBD, 3xHbic, PosIon, 2x Exvol	13.8	107.6	160.1	52.5	0.91	0.23	25.06	95%
9	8	Hbic, PosIon, 2xRingArom	17.9	117.1	188.3	71.2	0.94	0.17	17.45	95%
10	3	2xHBD, 2xHbic	16.8	112.5	188.3	75.8	0.96	0.08	7.41	95%
	7	HBD, 2xHbic, RingArom	16.8	114.9	188.3	73.5	0.95	0.05	4.21	95%
	10	2xHBD, Hbic, RingArom	16.8	116.6	188.3	71.7	0.93	0.09	8.31	95%
11	4	HBD, 3xHbic, PosIon.	15.2	118.6	188.3	69.7	0.92	0.23	25.48	95%
	6	2xHBD, 3xHbic	15.2	119.8	188.3	68.5	0.92	0.08	7.84	95%
	9	3xHbic, PosIon, RingArom	15.2	121.5	188.3	66.9	0.91	0.18	18.56	95%
12	4	2xHBD, 3xHbic, PosIon, RingArom	12.8	115.8	188.3	72.5	0.92	0.21	23.09	95%
	6	2xHBD, 2xHbic, PosIon	12.8	120	188.3	68.3	0.89	0.24	26.7	95%
	7	HBD, 3xHbic, RingArom	12.8	120	188.3	68.3	0.9	0.09	8.04	95%

Table 1.4: Success criteria of representative pharmacophoric hypotheses (cluster centers).

RUN[a]	Hypotheses[b]	Features	Cost				R^d	Global R^e	F-stat[f]	Cat. scramble
			Config.	Total	Null	Residual[c]				
	1	HBD, 3xHbic, 9xExVols	17.9	108.4	188.3	79.9	0.98	0.11	11.13	95%
13	9	HBA,2xHbic,PosIon, RingArom,5ExVols	17.9	116.1	188.3	72.2	0.94	0.18	18.35	95%
	1	2xHbic, PosIon, RingArom, 2x ExVol	17.9	109.2	188.3	79.1	0.97	0.17	17.84	95%
14	2	HBD, 2xHbic, PosIon, 6x ExVols	17.9	109.5	188.3	78.8	0.97	0.23	25.3	95%
	3	HBD, 2xHbic, PosIon, RingArom, 1x ExVols	15.2	113.2	188.3	75.1	0.94	0.14	14.19	95%
15	6	HBA, HBD, 2xHbic, PosIon.	15.2	115.3	188.3	73.1	0.93	0.22	24.9	95%
	7	HBA, HBD, Hbic, RingArom	18.0	96.2	134	37.8	0.87	0.09	8.14	95%
	9	HBD, 2xHbic, PosIon	18.0	98.2	134	35.8	0.83	0.2	21.83	95%
17	10	HBD, Hbic, 2xRingArom	18.0	98.3	134	35.7	0.83	0.06	5.93	95%
	4	2xHBD, 2xHbic	17.0	91.5	134	42.5	0.88	0.07	6.7	95%
18	7	HBD, Hbic, PosIon, RingArom	17.0	93.6	134	40.4	0.86	0.19	19.62	95%
	4	HBA, HBD, 2xHbic, PosIon	15.4	85.3	134	48.8	0.91	0.25	29.41	95%
	6	HBD, 2xHbic, PosIon, RingArom	15.4	88.3	134	45.7	0.9	0.21	23.26	95%
19	9	2xHBD, 2xHbic, PosIon	15.4	94.6	134	39.4	0.84	0.14	14.13	95%
21	7	HBA, 2xHBD, Hbic, 3x ExVols	18.0	89.5	134	44.5	0.91	0.04	4.02	95%
	7	2xHBD, 3xHbic, 2x ExVols	14.5	91.2	134	42.8	0.86	0.23	25.94	95%
23	8	HBD, 3x Hbic, RingArom , 6x ExVols	14.5	91.9	134	42.1	0.86	0.14	13.64	95%
	9	HBA, HBD, 3x Hbic, 3x ExVols	14.5	92.5	134	41.5	0.87	0.07	6.74	95%
	3	HBD, 3xHbic, PosIon 3xExVols	13.8	81.9	134	52.1	0.93	0.24	26.98	95%
24	6	2xHBD, 2xHbic, PosIon 3xExVols	13.8	87.8	134	46.2	0.88	0.3	36.33	95%
	2	2xHbic, PosIon, RingArom	16.7	111	188.3	77.3	0.96	0.17	18.06	95%
26	5	Hbic, PosIon, 2xRingArom	16.7	113	188.3	75.3	0.95	0.17	18.14	95%
	9	3xHbic, PosIon, RingArom	15.0	121	188.3	67.3	0.92	0.17	17.7	95%
27	10	2xHbic, PosIon, 2xRingArom	15.0	121.2	188.3	67.1	0.9	0.07	6.95	95%

Table 1.4: Success criteria of representative pharmacophoric hypotheses (cluster centers).

RUN[a]	Hypotheses[b]	Features	Cost				R^{d}	Global R^{e}	F-stat[f]	Cat. Scramble
			Config.	Total	Null	Residual[c]				
28	1	HBA, HBD, 3xHbic	12.9	107.5	188.3	80.9	0.97	0.07	6.05	95%
	5	HBD, 3xHbic, PosIon	12.9	116.9	188.3	71.4	0.91	0.23	25.76	95%
	6	2xHBD, 2xHbic, PosIon	12.9	120.2	188.3	68.1	0.89	0.19	19.89	95%
	8	HBD, 2xHbic, PosIon, RingArom	12.9	120.7	188.3	67.6	0.89	0.23	25.25	95%
	9	HBA, 3xHbic, PosIon	12.9	121.2	188.3	67.1	0.9	0.19	19.72	95%
29	5	HBA, 2xHbic, PosIon, RingArom, 5xExVols	17.8	115.5	188.3	72.8	0.94	0.18	18.62	95%
30	9	HBD, 2xHbic, PosIon, 3xExVols	16.7	114	188.3	74.3	0.95	0.22	24.52	95%
31	2	HBD, 2xHbic, PosIon, 5xExVols	15	108.9	188.3	79.4	0.97	0.25	29.27	95%
	8	**HBA, HBD, 2xHbic, PosIon, 5xExVols**	**15**	**115.2**	**188.3**	**73.1**	**0.93**	**0.25**	**29.01**	**90%**
	9	HBA, 2xHbic, PosIon, RingArom, 3xExVols	15	115.9	188.3	72.4	0.92	0.19	19.83	90%
	10	HBA, HBD, 2xHbic, PosIon, 4xExVols	15	115.9	188.3	72.4	0.93	0.06	5.98	90%
32	4	2xHBD, 3xHbic, 4xExVols	12.9	110.8	188.3	77.5	0.95	0.19	19.82	95%
33	6	HBA, 2xHbic, RingArom	14.8	103.8	131.9	28.1	0.83	0.05	4.69	95%
	7	Hbic, NegIon, 2xRingArom,	14.8	103.9	131.9	28	0.83	0.002	0.24	95%
	8	HBA, Hbic, 2xRingArom	14.8	103.9	131.9	28	0.83	0.02	2.2	95%
	10	HBA, 2xHbic, RingArom	14.8	104.3	131.9	27.6	0.82	0.05	4.15	95%
34	5	2xHbic, NegIon, RingArom	13.7	101.5	131.9	30.4	0.84	0.002	0.19	95%
	7	HBA, HBD, Hbic, RingArom	13.7	102.2	131.9	29.7	0.83	0.01	1.25	95%
36	5	HBA, 3xHbic, RingArom	10.1	100.5	131.9	31.4	0.81	0.04	3.44	95%
	8	HBD, 3xHbic, NegIon	10.1	101.9	131.9	30	0.82	0.04	3.71	95%
37	7	HBD, Hbic, NegIon, RingArom, 5xExVols	14.9	103.9	131.9	28	0.84	0.003	0.3	95%
38	9	HBA, 3xHbic, 1xExVol	13.8	103.5	131.9	28.4	0.82	0.05	4.53	95%
	10	HBA, 2xHbic, RingArom	13.8	103.7	131.9	28.2	0.82	0.05	4.14	95%
39	1	HBA, 3xHbic, 9xExVols	10.5	85.4	131.9	46.5	0.95	0.03	2.84	85%
	5	HBA, HBD, 3xHbic	10.5	98.6	131.9	33.3	0.83	0.05	4.43	95%
	7	**HBA, 4xHbic**	**10.5**	**99.2**	**131.9**	**32.8**	**0.83**	**0.01**	**0.58**	**95%**

Table 1.4: Success criteria of representative pharmacophoric hypotheses (cluster centers).

RUN[a]	Hypotheses[b]	Features	Cost				R^d	Global R^e	F-stat[f]	Cat. Scramble
			Config.	Total	Null	Residual[c]				
40	10	HBD, 2xHbic, NegIon, RingArom	10.1	104.3	131.9	27.6	0.8	0.002	0.2	95%
42	7	2xHBD, Hbic, RingArom	17.5	135.8	192.9	57.1	0.9	0.07	6.21	95%
	8	2xHBD, 2xHbic	17.5	136	192.9	57	0.9	0.11	10.09	95%
	9	HBD, 2xHbic, RingArom	17.5	136.1	192.9	56.8	0.89	0.06	5.38	95%
44	5	HBA, HBD, 3xHbic	13.7	141.3	192.9	51.6	0.85	0.1	9.95	95%
	6	2xHBD, 3xHbic	13.7	141.5	192.9	51.4	0.84	0.13	12.36	95%
45	2	HBA, 2xHBD, Hbic, 5xExVols	18.5	133.1	192.9	59.8	0.92	0.07	6.6	95%
	4	HBA, HBD, 2xHbic	18.5	137.6	192.9	55.4	0.89	0.11	10.31	95%
	7	2xHBD, 2xHbic, 1xExVol	18.5	138.5	192.9	54.4	0.88	0.07	6.27	95%
	9	2xHBD, 2xHbic, 5xExVols	18.5	139.2	192.9	53.7	0.89	0.13	12.92	95%
	10	HBD, 2xHbic, RingArom, 3xExVols	18.5	139.9	192.9	53	0.88	0.06	5.35	95%
46	1	HBD, 2xHbic, RingArom, 3xExVols	17.5	128	192.9	64.9	0.94	0.08	7.62	95%
	7	2xHBD, Hbic, RingArom	17.5	136	192.9	56.9	0.9	0.05	4.49	95%
47	1	2xHBA, 3xHbic, 9xExVols	15.3	123.4	192.9	69.5	0.96	0.13	12.57	90%
	9	**HBD, 3xHbic, RingArom**	**15.3**	**143.9**	**192.9**	**49**	**0.83**	**0.11**	**10.46**	**95%**
48	3	HBA, 3xHbic, RingArom, 5xExVols	13.7	123.8	192.9	69.1	0.95	0.14	14.53	95%
	8	HBA, HBD, 3xHbic	13.7	135.1	192.9	57.8	0.88	0.14	13.82	95%
49	4	2xHBD, 2xHbic	17.4	129.3	199.7	70.4	0.93	0.06	5.7	95%
	10	2xHBD, 2xHbic	17.4	135.9	199.7	63.8	0.89	0.03	2.47	95%
50	5	2xHBD, Hbic, RingArom	16.3	129.2	199.7	70.5	0.91	0.07	6.46	95%
	6	2xHBD, Hbic, RingArom	16.3	130	199.7	69.7	0.9	0.08	7.53	95%
	8	2xHBD, 2xHbic	16.3	130.4	199.7	69.3	0.91	0.07	6.7	95%
53	3	HBD, 2xHbic, RingArom, 3xExVols	17.4	127.4	199.7	72.3	0.92	0.06	5.55	95%
	5	2xHBD, 2xHbic	17.4	132.4	199.7	67.3	0.89	0.07	6.66	95%
	7	HBD, 2xHbic, RingArom, 5xExVols	17.4	133.1	199.7	66.6	0.88	0.06	5.62	95%
	9	2xHBD, 2xHbic	17.4	134.7	199.7	65	0.88	0.07	6.12	95%

Table 1.4: Success criteria of representative pharmacophoric hypotheses (cluster centers).

RUN[a]	Hypotheses[b]	Features	Cost				R^d	Global	F-	Cat.
			Config.	Total	Null	Residual[c]		R^e	stat[f]	Scramble
	2	2xHBD, 2xHbic, 1xExVol	16.3	118.1	199.7	81.6	0.96	0.07	6.02	95%
54	8	2xHBD, 2xHbic, 1xExVol	16.3	127.8	199.7	71.9	0.92	0.08	7.37	95%
	9	2xHBD, Hbic, RingArom, 1xExVol	16.3	127.8	199.7	71.9	0.91	0.07	6.06	95%
55	1	2xHBD, 3xHbic, 5xExVols	14.2	116.7	199.7	83.0	0.96	0.12	12.2	95%
	2	2xHBD, 3xHbic, 4xExVols	14.2	127.5	199.7	72.2	0.92	0.11	10.37	95%
56	2	HBD, 3xHbic, RingArom, 4xExVols	11.8	124.2	199.7	75.5	0.91	0.13	12.46	95%
	6	2xHBD, 3xHbic, 1xExVol	11.8	129	199.7	70.7	0.88	0.08	7.82	95%
57	6	2xHBD, Hbic, RingArom	17.2	131.7	199.7	68.0	0.9	0.07	6.26	95%
	7	2xHBD, 2xHbic,	17.2	133.4	199.7	66.3	0.9	0.08	7.23	95%
	8	2xHBD, 2xHbic	17.2	133.5	199.7	66.2	0.89	0.07	6.48	95%
	10	HBA, HBD, 2xHbic	17.2	135.7	199.7	64.0	0.89	0.06	5.36	95%
58	8	HBD, Hbic, 2xRingArom	16.2	130.7	199.7	69.0	0.89	0.09	8.13	95%
	10	2xHBD, Hbic, RingArom	16.2	132.3	199.7	67.4	0.91	0.04	3.13	95%
59	4	HBD, 3xHbic, RingArom, 3xExVols	14	129.2	199.7	70.5	0.89	0.17	17.89	95%
	9	3xHbic, 2xRingArom	14	136.9	199.7	62.8	0.87	0.03	2.88	95%
	10	HBA, HBD, 3xHbic, 1xExVol	14	136.9	199.7	62.8	0.85	0.08	7.96	95%
60	4	2xHBD, 3xHbic, 2xExVols	12	127.8	199.7	71.9	0.89	0.14	14.23	95%
61	3	2xHBD, 2xHbic, 5xExVols	17.2	125.2	199.7	74.5	0.95	0.1	9.15	95%
61	6	2xHBD, 2xHbic, 4xExVols	17.2	129.1	199.7	70.6	0.92	0.06	5.79	95%
61	7	HBD, 3xHbic, 6xExVols	17.2	129.1	199.7	70.6	0.91	0.001	0.11	95%
61	9	2xHBD, 2xHbic, 2xExVols	17.2	129.4	199.7	70.4	0.92	0.06	5.21	95%
61	10	HBA, HBD, Hbic, 1xExVol	17.2	130.1	199.7	69.6	0.9	0.04	3.27	95%
63	6	HBA, HBD, 3xHbic, 5xExVols	14.8	132.1	199.7	67.6	0.89	0.11	11.12	95%
65	3	2xHBD, Hbic, RingArom	18.5	101.7	162.4	60.7	0.98	0.07	6.71	95%
	6	HBA, HBD, 2xHbic, PosIon	18.5	102.9	162.4	59.5	0.98	0.2	21.42	95%

Table 1.4: Success criteria of representative pharmacophoric hypotheses (cluster centers).

RUN[a]	Hypotheses[b]	Features	Cost				R^{d}	Global R^{e}	F-stat[f]	Cat. Scramble
			Config.	Total	Null	Residual[c]				
66	9	HBD, PosIon, 2x RingArom	17.9	101.7	162.4	60.7	0.97	0.16	16.84	95%
67	2	3xHbic, PosIon, RingArom	16.4	100.2	162.4	62.2	0.97	0.09	8.82	95%
	10	2xHBA, 2xHbic, PosIon	16.4	101.2	162.4	61.2	0.96	0.17	17.54	95%
68	4	HBA, 3xHbic, PosIon	14.3	98.5	162.4	64	0.97	0.16	16.3	95%
	6	HBD, 3xHbic	14.3	99.2	162.4	63.2	0.96	0.11	11.07	95%
	8	HBD, 2xHbic, PosIon, RingArom	14.3	99.4	162.4	63	0.96	0.2	21.2	95%
	8	HBD, 2xHbic, PosIon, RingArom	14.3	99.4	162.4	63	0.96	0.16	16.84	95%

[a]Correspond to runs in Table 1.3 under supplementary materials, [b]High ranking representative hypotheses (in their corresponding clusters, see section **1.3.1.2.5 Clustering of the Generated Pharmacophore Hypotheses**. [c]Difference between total cost and the cost of the corresponding null hypotheses. [d]Correlation coefficients between pharmacophore-based bioactivity estimates (calculated from equation (1.3) and bioactivities of corresponding training compound (subsets in table B under supplementary material). [e]Correlation coefficients between pharmacophore-based bioactivity estimates and bioactivities of all collected compounds. [f]Fisher statistic calculated based on the linear regression between the fit values of all collected inhibitors (**1-88**, table 1.1) against pharmacophore hypothesis (employing the "best fit" option and equation (1.4)) and their respective anti- CaMKIIδ bioactivities (log ($1/IC_{50}$) values). [g]ranking of hypotheses is as generated by CATALYST in each automatic run. [h]Bolded pharmacophores appeared in the best QSAR equations.

1.4.3 QSAR Modeling

Despite that pharmacophoric binding hypotheses provide excellent insights into ligand-macromolecule recognition and can be used to mine for new biologically interesting scaffolds, their predictive value as 3D-QSAR models is limited by steric shielding and bioactivity-enhancing or reducing auxiliary groups (Taha *et al.*, 2008, a,b&c; Taha *et al.*, 2007; Taha *et al.*, 2010; Al-masri *et al*; 2008; Al-Nadaf *et al*; 2010; Abu-Hammad *et al*, 2009; Abu Khalaf *et al*, 2010; Al-Sha'er *et al*, 2010, a&b; Abdula *et al*, 2011; Abu Khalaf *et al*, 2011). This point combined with the fact that pharmacophore modeling of CaMKII□ inhibitors furnished several binding hypotheses of comparable success criteria prompted us to employ classical QSAR analysis to search for the best combination of pharmacophore(s) and other 2D descriptors capable of explaining bioactivity variation across the whole list of collected inhibitors (**1-88**, table 1.1 and figure 1.3). We employed genetic function approximation and multiple linear regression QSAR (GFA-MLR-QSAR) analysis to search for an optimal QSAR equation(s).

The fit values obtained by mapping representative hypotheses (137 models) against collected CaMKII□ inhibitors (**1-88**, table 1.1 and figure 1.1) were enrolled, together with other physicochemical descriptors (around 250), as independent variables (genes) in GFA-MLR-QSAR analysis (see section **1.3.1.3 QSAR Modeling** in **Experimental**) (Taha *et al.*, 2008, a,b&c; Taha *et al.*, 2007; Taha *et al.*, 2010; Al-masri *et al*; 2008; Al-Nadaf *et al*; 2010; Abu-Hammad *et al*, 2009; Abu Khalaf *et al*, 2010; Al-Sha'er *et al*, 2010, a&b; Abdula *et al*, 2011; Abu Khalaf *et al*, 2011; CERIUS2 QSAR Users' Manual, 2005; Ramsey and Schafer, 1997). However, since it is essential to access the predictive power of the resulting QSAR models on an external set of inhibitors, we randomly selected 16 molecules (marked with double

asterisks in table 1.1, see section **1.3.1.3 QSAR Modeling** in **Experimental**) and employed them as external test molecules for validating the QSAR models (r^2_{PRESS}). Moreover, all QSAR models were cross-validated automatically using the leave-one-out cross-validation in CERIUS2 (CERIUS2 QSAR Users' Manual, 2005; Ramsey and Schafer, 1997).

Equation (1.10) shows the details of the optimal QSAR model. Figure 1.3 shows the corresponding scatter plots of experimental versus estimated bioactivities for the training and testing inhibitors.

$$\begin{aligned} Log(1/IC_{50}) = \; & 3.35 + 17.5x10^{-2} (Hypo8/31) + 9.67x10^{-2} (Hypo9/47) + 10.8x10^{-2} (Hypo7/39) \\ & - 9.5x10^{-2}(SdO) - 3.73(JursFPSA1) + 5.41(JursRPSA) - 0.80(AtypeC28) \\ & - 8.0x10^{-2}(SaaCH) \end{aligned}$$

$$r^2_{72} = 0.70, \text{ F-statistic} = 18.19 \quad r^2_{BS} = 0.59, \; r^2_{LOO} = 0.71, \; r^2_{PRESS(16)} = 0.600 \; \ldots\ldots\ldots\ldots(1.10)$$

where, r^2_{72} is the correlation coefficient against 72 training compounds, r^2_{LOO} is the leave-one-out correlation coefficient, r^2_{BS} is the bootstrapping regression coefficient and r^2_{PRESS} is the predictive r^2 determined for the 16 test compounds (CERIUS2 QSAR Users' Manual, 2005; Ramsey and Schafer, 1997). Hypo7/39, Hypo8/31 and Hypo9/47 represent the fit values of the training compounds against these three pharmacophores (bolded models in table 1.4 and figures 1.4, 1.5, and 1.6) as calculated from equation (1.4) under section. **1.3.1.2.4 Assessment of the Generated Hypotheses**.

JursFPSA1 is the fractional positively charged partial surface area (calculated by dividing sum of the solvent-accessible surface areas of all positively charged atoms by

the total molecular solvent accessible surface area), JursRPSA is the relative polar surface area calculated by dividing total polar surface area by the total molecular solvent-accessible surface area, AtypeC28 is one of the thermodynamic AlogP_atypes family of descriptors and it encodes for the hydrophobic contributions of certain carbon atom types in LogP, SaaCH is an electrotopological state descriptor for aromatic CH groups, and SdO is the electrotopological state index of carbonyl fragments (CERIUS2 QSAR Users' Manual, 2005) (See table 4.2 in Appendix).

Emergence of three orthogonal pharmacophoric models, i.e., Hypo7/39, Hypo8/31 and Hypo9/47, of cross-correlation $r^2 \leq 0.21$ (Table 1.5) in equation (1.10) suggests they represent three complementary binding modes accessible to ligands within the binding pocket of CaMKII, i.e., a pharmacophore can optimally explain the bioactivities of some training inhibitors, while the others explain the remaining inhibitors. Similar conclusions were reached about the binding pockets of other targets based on QSAR analysis (Taha *et al.*, 2008, a,b&c; Taha *et al.*, 2007; Taha *et al.*, 2010; Al-masri *et al*; 2008; Al-Nadaf *et al*; 2010; Abu-Hammad *et al*, 2009; Abu Khalaf *et al*, 2010; Al-Sha'er *et al*, 2010, a&b; Abdula *et al*, 2011; Abu Khalaf *et al*, 2011).

Table 1.5: The cross correlation between the successful hypothesis in Equation 1.10.

	Hypo8/31	Hypo 9/47	Hypo 7/39
Hypo8/31	1		
Hypo 9/47	0.10	1	
Hypo 7/39	0.21	0.05	1

Figures 1.4C and 1.5C show how Hypo8/31 and Hypo9/47 map **85** (IC_{50} = 0.009 μM), while figure 1.7C shows how Hypo7/39 maps **68** (IC_{50} = 0.21 μM). The X, Y, and Z coordinates of the three pharmacophores are shown in table 1.6.

Emergence of JursFPSA1 in equation (1.10) in association with significant negative contribution suggests that molecules of diffuse positive charges, i.e., over wide surface areas (larger atoms or fragments), disfavor binding to CaMKII compared to molecules of focused positive charges, i.e., on smaller-sized fragments, as implied from the emergence of Hypo8/31 (includes a PosIon feature) in equation (1). On the other hand, the combined emergence of JursRPSA and SdO, in association with positive and negative regression coefficients, respectively, suggests that compounds of surface-disseminated polarity (of higher JursRPSA values) prefer binding compared to compounds of focused polar areas. For example, carbonyl fragments (e.g., as carboxyl or nitro groups) promote bias towards SdO, and therefore, negatively influences affinity.

The interesting combination of descriptors in equation (1.10) suggests that CaMKII binding pocket is predominantly polar, which explains the superior affinities of ligands having diffuse polarities. This conclusion agrees with negative contributions of the hydrophobic descriptors AtypeC28 and SaaCH in equation (1.10). However, focused polarities (e.g., carbonyl groups) seem to disfavor binding probably due to hydration that competes with binding. On the other hand, the negative effect of diffuse positive charges on affinity suggests that CaMKII binding pocket contains positively-charged residues that repel ligands of diffuse positive charges e.g compound **75** (Table 1.1), while the presence of correctly positioned focused positive groups promote binding through electrostatic attraction with tightly-positioned corresponding negative residue in the binding pocket.

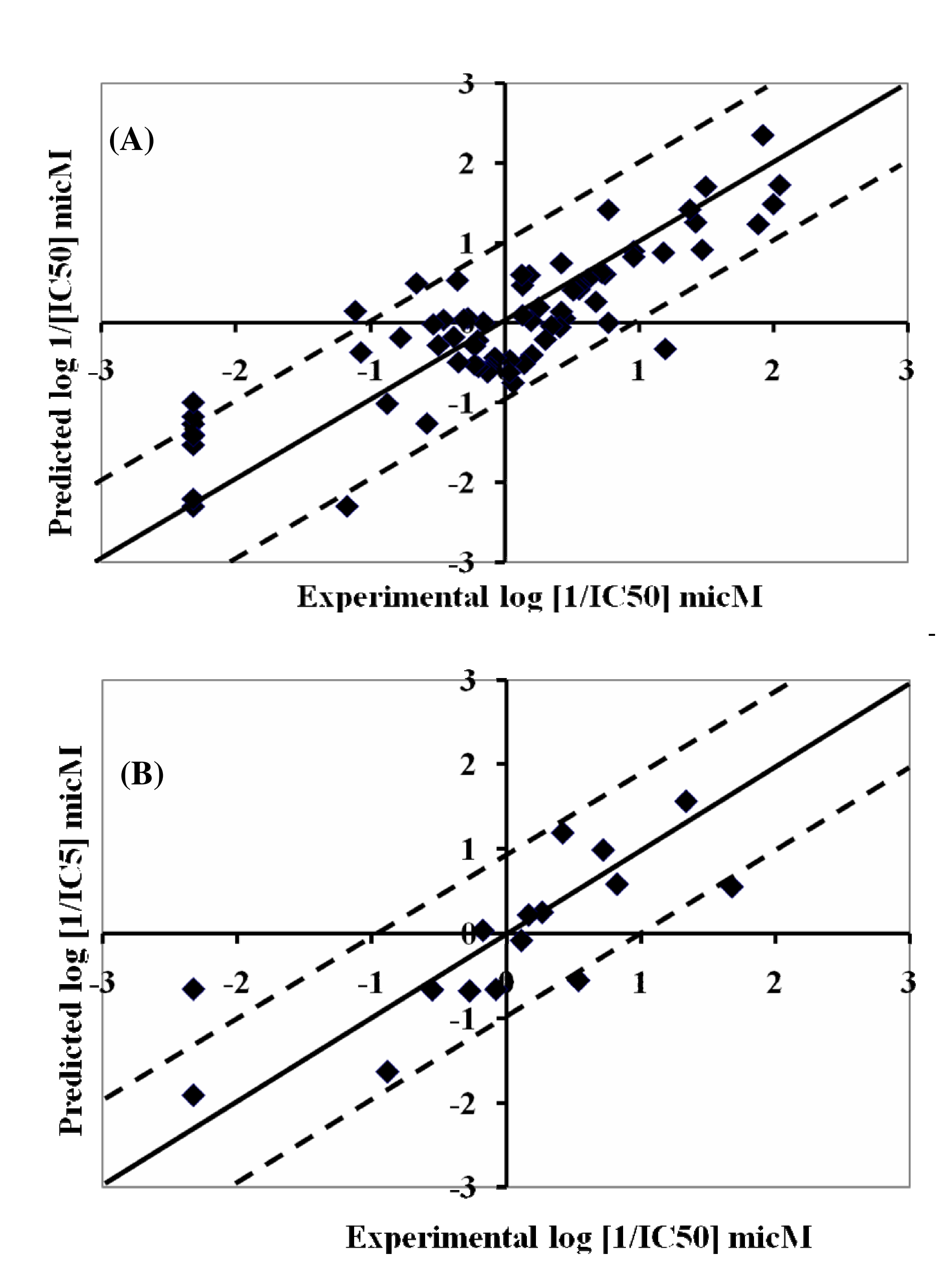


Figure 1.3: Experimental versus **(A)** fitted (72 compounds, $r2_{LOO}$= 0.71), and **(B)** predicted (16 compounds, $r2_{PRESS}$= 0.60) bioactivities calculated from the best QSAR model equation (1.10). The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively, whereas the dotted lines indicate 1.0 log point error margins

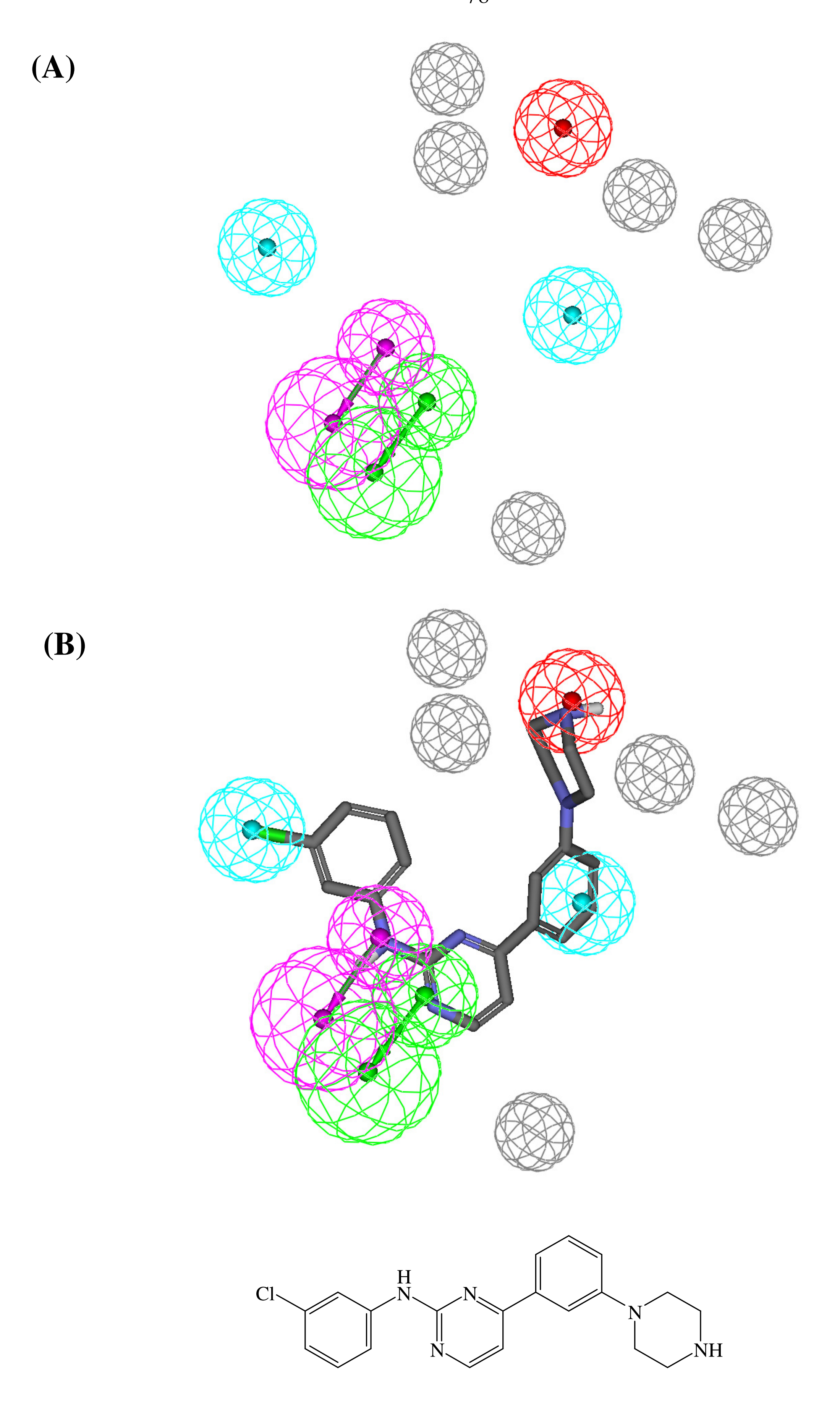


Figure 1.4: (**A**) Pharmacophoric features of Hypo8/31: HBD as pink vectored spheres, HBA as green vectored spheres, Hbic as blue spheres, PosIon as red spheres, Exclusion volumes as grey spheres, **(B)** Hypo8/31 fitted against **85** (IC_{50} = 0.009 μM, table 1.1), **(C)** Chemical structure of **85**.

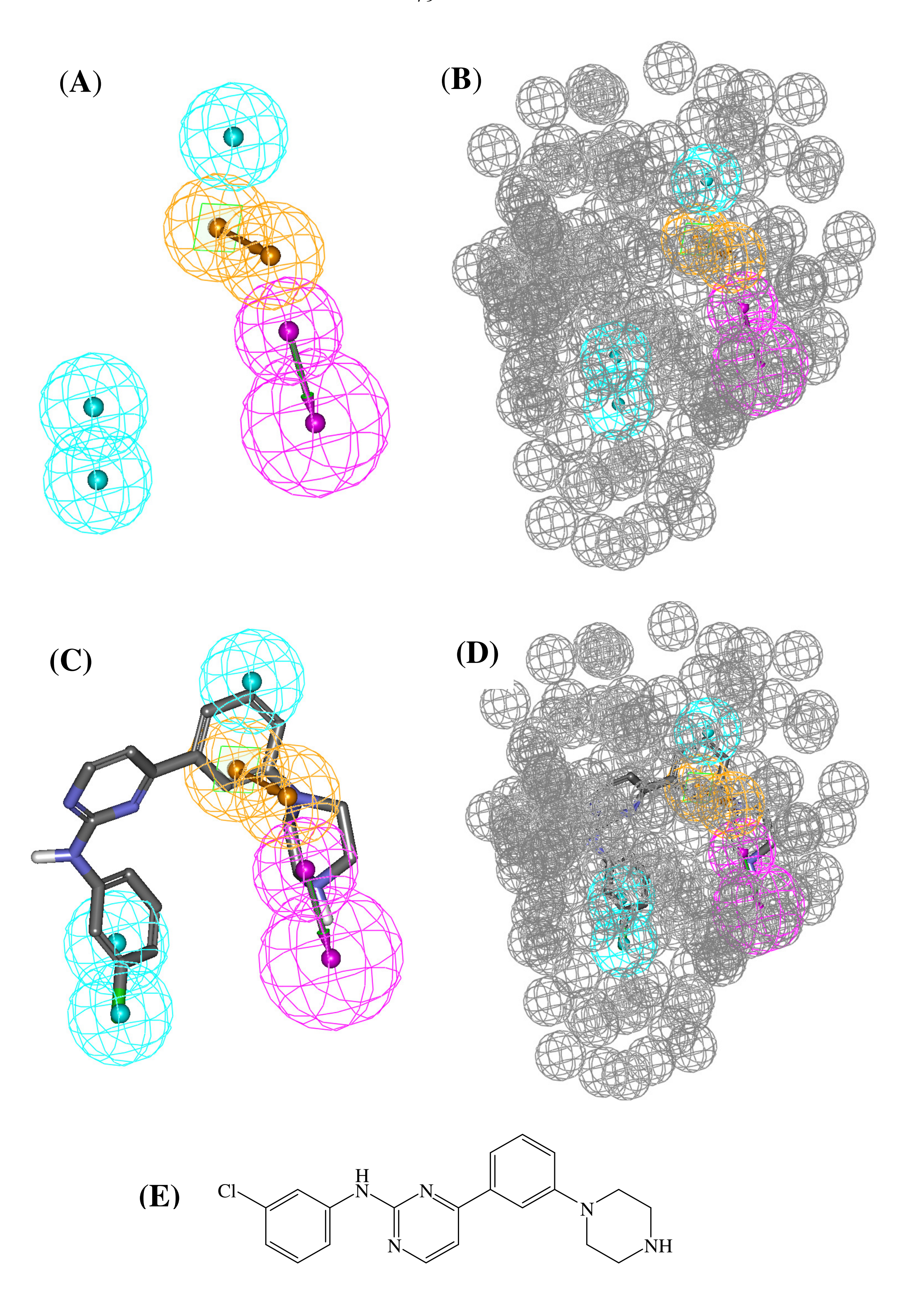


Figure 1.5: **(A)** and **(B)** show Pharmacophoric features of Hypo 9/47 without and with steric-refinement, respectively. HBD as pink vectored spheres, Hbic as blue spheres, RingArom as vectored orange spheres, Exclusion volumes as grey spheres, **(C)** and **(D)** Hypo 9/47 mapped against **85** (IC_{50} = 0.009 μM, table 1.1), **(E)** Chemical structure of **85**.

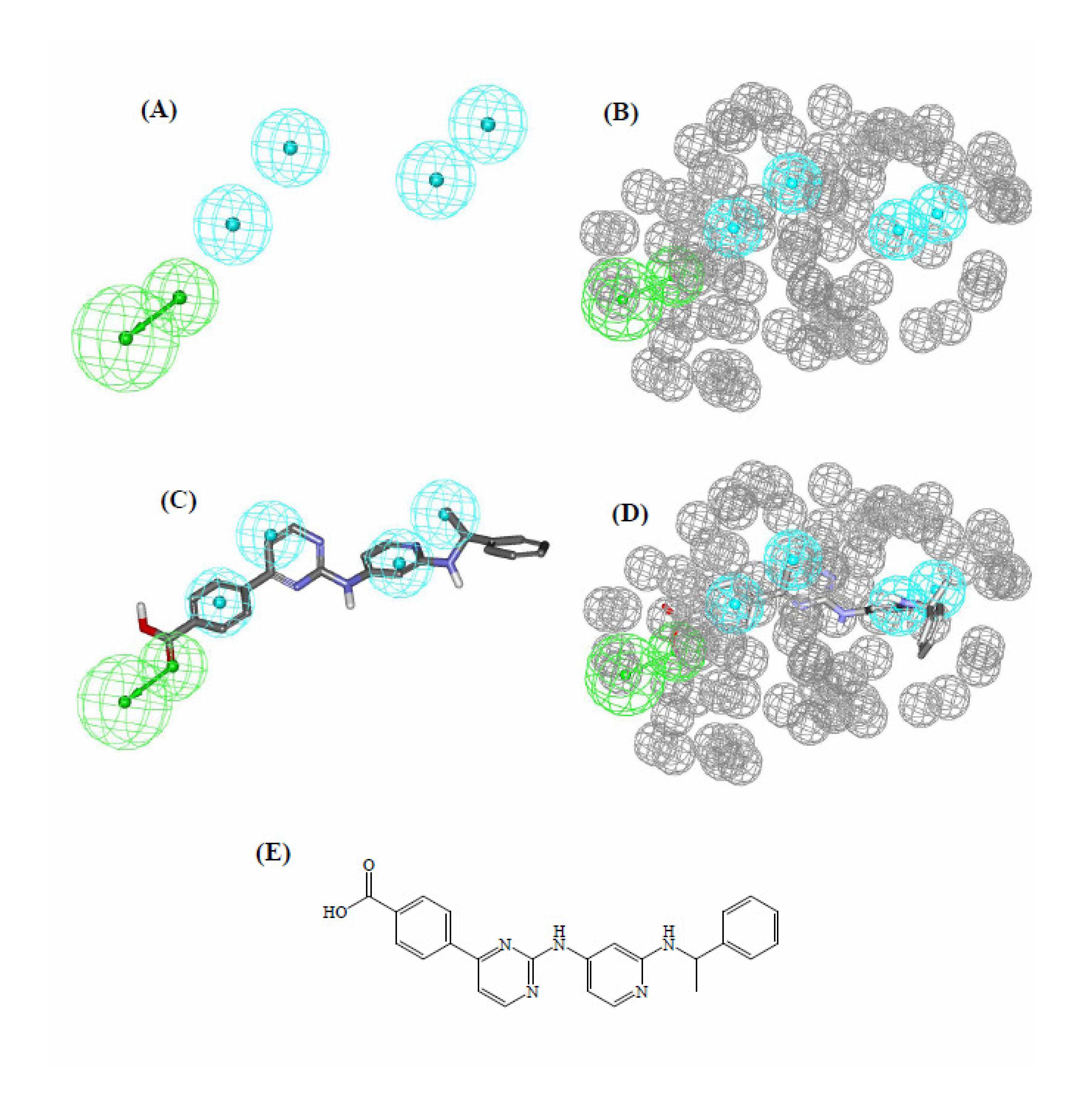


Figure 1.6: **(A)** and **(B)** show Pharmacophoric features of Hypo7/39 without and with steric-refinement, respectively. HBA as green vectored spheres, Hbic as blue spheres, Exclusion volumes as grey spheres, **(C)** and **(D)** Hypo7/39 mapped against **68** (IC_{50} = 0.21 μM, table 1.1) without and with steric-refinement, respectively, **(E)** Chemical structure of **68**.

Table 1.6: Pharmacophoric features and corresponding weights, tolerances and 3D coordinates of Hypo8/31, Hypo9/47 and Hypo7/39.

Model	definitions		Chemical Features							
			HBA[e]		**HBD[f]**		**Hbic[g]**	**Hbic**	**PosIons[h]**	
Hypo8/31[a,j]	Weights		2.48		2.48		2.48	2.48	2.48	
	Tolerances[d]		1.60	2.20	1.60	2.20	1.60	1.60	1.60	
	Coordinates	X	-1.93	-4.11	-3.06	-4.84	-4.61	3.58	2.37	
		Y	-0.94	-2.03	0.83	-0.15	5.68	-0.48	3.05	
		Z	-3.02	-4.83	-2.13	-4.34	-2.20	-0.38	4.93	
			HBD		**Hbic**		**Hbic**	**Hbic**	**RingArom[i]**	
Hypo9/47[b]	Weights		1.86		1.86		1.86	1.86	1.86	
	Tolerances		1.60	2.20	1.60		1.60	1.60	1.60	1.60
	Coordinates	X	-0.93	0.72	-2.98		3.80	2.66	-1.34	-3.67
		Y	-2.66	-3.66	-1.38		1.64	2.80	-1.02	-1.14
		Z	-2.04	-4.36	3.40		-2.12	-4.82	1.22	-0.66
			HBA		**Hbic**		**Hbic**	**Hbic**	**Hbic**	
Hypo7/39[c]	Weights		2.04	2.04	2.04		2.04	2.04	2.04	
	Tolerances		1.60	2.20	1.60		1.60	1.60	1.60	
	Coordinates	X	-5.21	-7.42	5.80		-1.04	4.92	-3.00	
		Y	-4.64	-6.68	4.14		2.14	1.78	-1.30	
		Z	-0.18	-0.18	3.46		-0.28	1.28	-0.64	

[a]Hypo8/31: the 8th pharmacophore hypothesis generated in the 31st HYPOGEN run (table 1.3 and table 1.4).

[b]Hypo9/47: the 9th pharmacophore hypothesis generated in the 47th HYPOGEN run (table 1.3 and table 1.4).

[c]Hypo7/39: the 7th pharmacophore hypothesis generated in the 39th HYPOGEN run (table 1.3 and table 1.4).

[d]Tolerances: refer to the radius of feature spheres

[e]HBA: Hydrogen Bond Acceptor feature.

[f]HBD: Hydrogen Bond Donor feature.

[g]Hbic: Hydrophobic feature.

[h]PosIons: Positive Ionizable feature.

[i]RingArom: Ring Aromatic feature.

[j]Hypo8/31 includes 5 exclusion spheres, each of 1.2 Å tolerance, at the following X,Y,Z coordinates: (6.782,-1.976,4.383), (-0.114,-5.669,-4.343), (-1.033,5.701,5.06), (-5.71,1.332, 6.713), (5.458,1.028,3.346).

1.4.4 Receiver Operating Characteristic (ROC) Curve Analysis and shape constrains.

To further validate the resulting models (both QSAR and pharmacophores), we subjected our QSAR-selected pharmacophores to receiver-operating characteristic (ROC) curve analysis. In ROC analysis, the ability of a particular pharmacophore model to correctly classify a list of compounds as actives or inactives is indicated by the area under the curve (AUC) of the corresponding ROC as well as other parameters, namely, overall accuracy, overall specificity, overall true positive rate and overall false negative rate (see 1.3.1.5 ROC Curve Analysis under Experimental for more details) (Verdonk *et al*, 2004; Kirchmair *et al*, 2008; Irwin *et al,* 2005; Triballeau *et al*, 2005).

Table 4 and figure 6 show the ROC results of QSAR-selected pharmacophores. Hypo9/47 and Hypo7/39 illustrated good overall performances with AUC values of 77.8%.and 84.3%, respectively. On the other hand, Hypo8/31 exhibited excellent performance with AUC value of 98.1%. This is not unexpected, as presence of positive ionizable, HBD and HBA features in Hypo8/31 should enhance its selectivity compared to Hypo9/47 and Hypo7/39.

The fair overall performances of Hypo9/47 and Hypo7/39 suggest certain degree of promiscuousity (Abu-Hammad et al, 2009). Therefore, we decided to complement the two pharmacophores with exclusion spheres employing HipHop-Refine module of CATALYST (CATALYST 4.11 Users' Manual, **2005**). Excluded volumes resemble sterically inaccessible regions within the binding site (see section 1.3.1.4 Addition of Exclusion Volumes under Experimental) (Clement *et al*, 2000). Accordingly, a

structurally diverse training subset was selected for addition of exclusion spheres to Hypo9/47 and Hypo7/39 (table 1.7). The training compounds were selected in such away that the bioactivities of weakly active compounds are explainable by steric clashes within the binding pocket. Figures 1.54B, 1.6B show sterically refined versions of Hypo9/47 and Hypo7/39, respectively.

Table 1.7 Training subset used for adding excluded spheres for Hypo9/47 and Hypo 7/39, using HIPHOP-REFINE module of CATALYST.

Compound[a]	IC_{50} (µM)	Principal value	MaxOmitFeat[b]	Ref[c]
4	210	0	0	1
8	210	0	2	1
22	0.034	2	0	1
23	2.32	0	2	1
24	15.02	0	2	1
28	210	0	2	2
33	210	0	2	2
39	1.87	0	2	2
41	0.74	1	0	3
51	0.01	2	0	3
52	0.013	2	0	3
53	0.038	2	0	3
54	0.032	2	0	3
58	0.012	2	0	3
65	0.92	1	0	4
67	1.96	1	0	4
71	2.22	1	0	4
74	0.17	2	0	4
77	0.064	2	0	4
79	210	0	0	4
80	210	0	1	4
81	210	0	1	4
85	0.009	2	0	4

[a]Compounds' numbers are as table 1.1
[b]MaxOmitFeat: Maximum omitted features.
[c]Ref: References: 1: Levy *et al*, a , 2008 ; 2: Levy *et al*, b , 2008, 3: Lu *et al*, 2008 , 4 :Mavunkel *et al*, 2008.

.

Clearly, from the figure 1.7 and table 1.8 the classification performances of sterically-refined versions of Hypo9/47 and Hypo7/39 improved significantly (compared to their unrefined versions) as reflected by their ROC AUC values, which shifted from 77.8% to 97.5% and from 84.3% to 94.6%, respectively.

Table 1.8: ROC curve analysis criteria for QSAR-selected pharmacophores and their sterically-refined versions.

Pharmacophore Model	ROC[a]-AUC[b]	ACC[c]	SPC[d]	TPR[e]	FNR[f]
Hypo8/31	98.1	97.1	98.6	46.2	1.4
Hypo 9/47	77.8	97.1	97.5	84.6	2.5
Hypo 7/39	84.3	97.1	99.3	23.1	7
Refined Hypo 9/47	97.5	97.1	99.5	18.0	5
Refined Hypo 7/39	94.6	97.2	99.5	15.8	5

[a]**ROC**: receiver operating characteristic curve.
[b]**AUC**: area under the curve.
[c]**ACC**: overall accuracy.
[d]**SPC**: overall specificity.
[e]**TPR**: overall true positive rate.
[f]**FNR**: overall false negative rate.

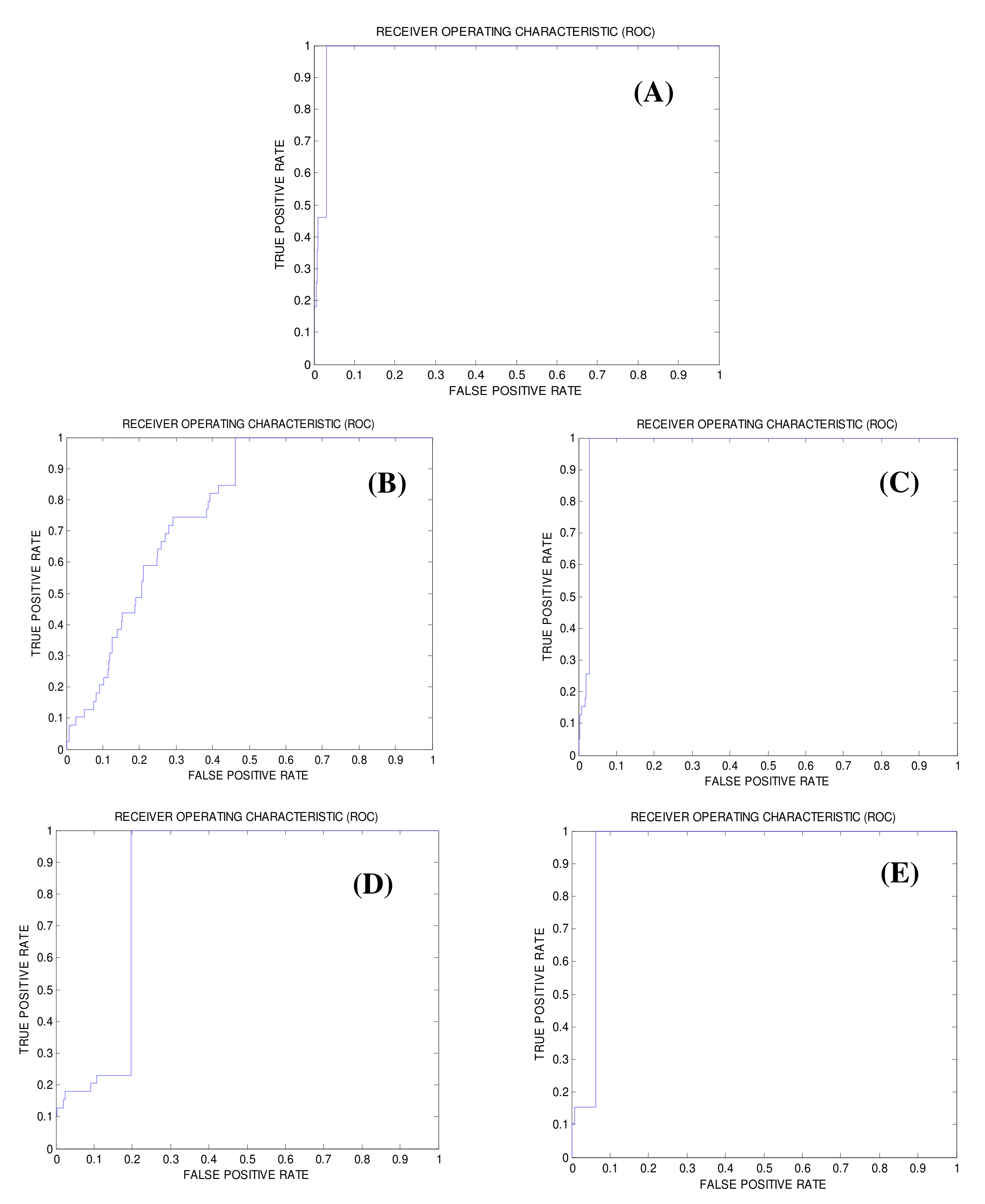


Figure 1.7: ROC curves of: **(A)** Hypo8/31, **(B)** Hypo9/47, **(C)** Sterically-Refined Hypo9/47, **(D)** Hypo7/39, **(E)** Sterically-Refined Hypo7/39**.**

1.4.5 *In-Silico* screening and Subsequent *In vitro* Evaluation

Hypo8/31, sterically-refined Hypo9/47 and sterically-refined Hypo7/39 were employed as 3D search queries against the NCI (238,819 structures) using the "Best Flexible Database Search" option implemented within CATALYST. Compounds that have their chemical groups spatially overlap (map) with corresponding features of the particular pharmacophoric model were captured as hits. Table 1.9 summarizes the numbers of captured hits by these pharmacophores. The hits were filtered based on Lipinski's and Veber's rules (Lipinski *et al.,* 2001; Veber *et al*., 2002). Enforcing drug-likeness pre-filters should help in finding hits more amenable for subsequent optimization into leads.

Table 1.9: Numbers of captured hits by sterically-refined versions of Hypo8/31, sterically-refined Hypo9/47 and Hypo7/39

		Pharmacophore models		
3D Database	**Post screening filtering[b]**	**Hypo8/31**	**Sterically-Refined Hypo9/47**	**Sterically-Refined Hypo7/39**
NCI[a]	Before	1823	13165	10697
	After	340	10477	3385

[a]NCI: national cancer institute list of available compounds (238,819 structures),[b]Using Lipinski's and Veber's rules.

Surviving hits were fitted against Hypo8/31, Hypo9/47 and Hypo7/39 (without exclusion volumes) and their fit values, together with other relevant molecular descriptors, were substituted in QSAR equation (1.10) to predict their anti-CaMKII IC_{50} values. The highest-ranking 60 available hits were evaluated *in vitro* against human recombinant CaMKII assay kit (Cyclex, Japan). Initially, the acquired hits were screened at 10 µM concentrations, subsequently; compounds of anti-CaMKII inhibitory percentages ≥ 50% at 10 µM were further assessed to determine their IC_{50} values. Although 37 exhibited measurable anti-CaMKII bioactivities, only 10 hits exhibited ≥ 50% inhibition at 10 µM and were, therefore, evaluated to determine their IC_{50} values. Figure 1.10 illustrates the dose/response plots of the active hits and standard inhibitor KN-62 under our bioassay conditions.

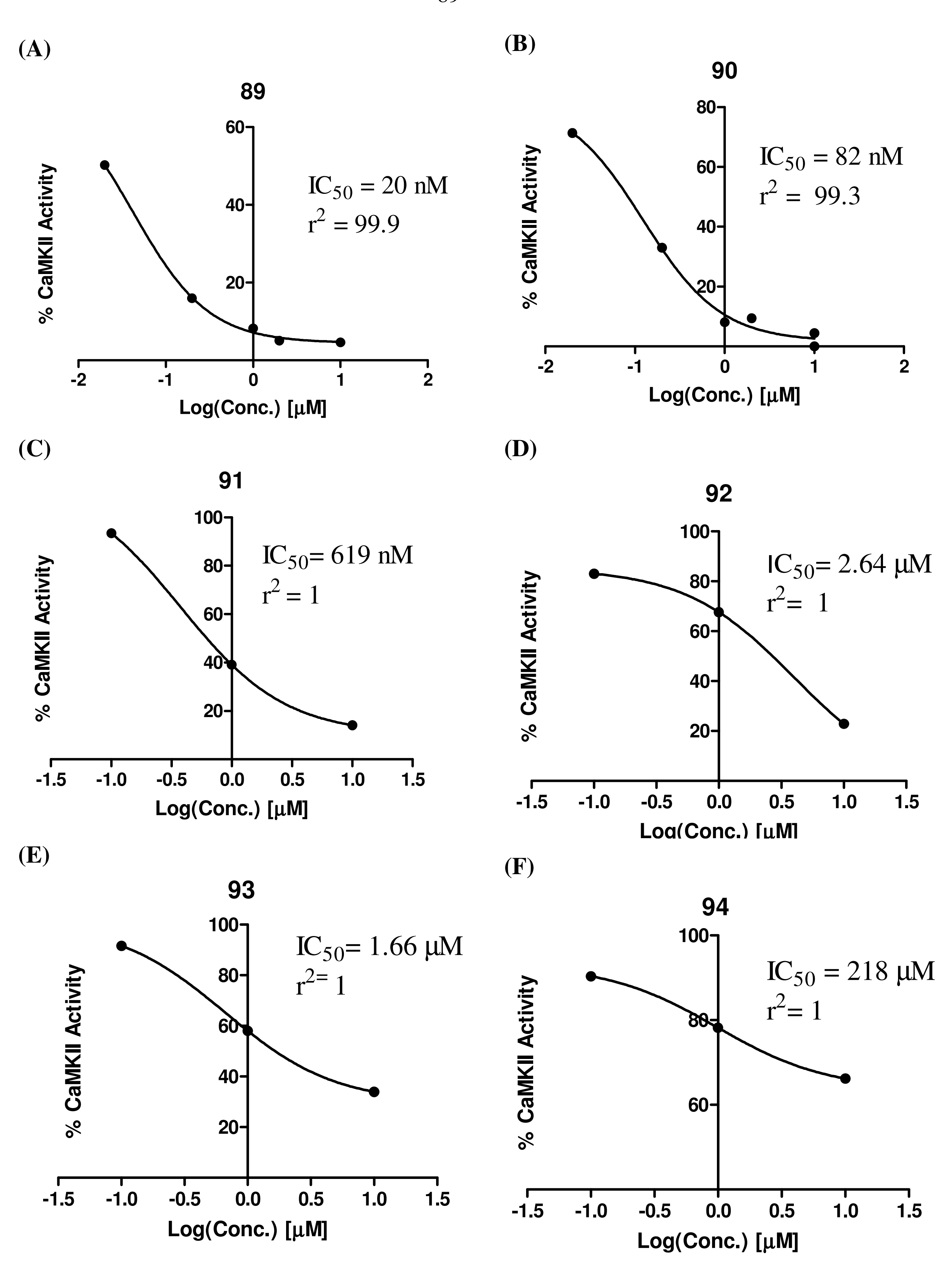


Figure 1.8: Dose-response plots of active hits: **89**, **90**, **91**, **92**, **93**, and **94**. Data fitting was performed using GraphPad Prism (version 5.04) via fitting against sigmoidal dose-inhibition model.

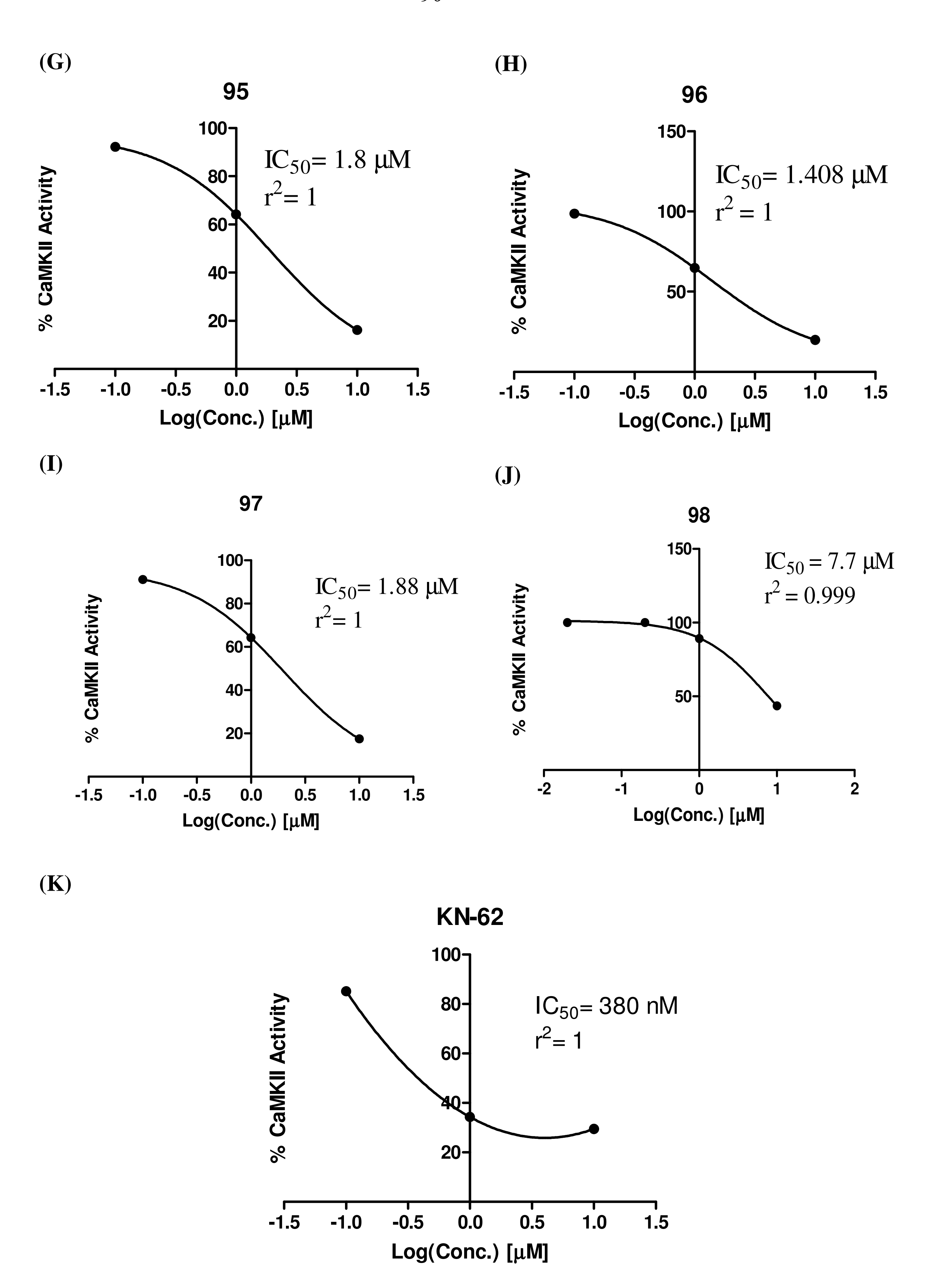


Figure 1.9: Dose-response plots of active hits: **95**, **96**, **97, 98, and standard inhibitor KN-62** . Data fitting was performed using GraphPad Prism (version 5.04) via fitting against sigmoidal dose-inhibition model.

Table 1.10 shows the most active hits (**89-98**), their corresponding fit values against Hypo8/31, Hypo9/47, and Hypo7/39, predicted bioactivities based on QSAR equation (1.10), and their *in vitro* bioactivities. Figures 1.10 and 1.11 show hits **89** and **92** and how they map corresponding QSAR-selected pharmacophores.

The bioassay procedure was validated by assessing the IC_{50} value of standard inhibitor KN-62 (Davies *et al*, 2000) via the same assay conditions. It exhibited IC_{50} value of 380 nM, which is rather comparable with the reported literature value (500 nM) (Davies *et al*, 2000).

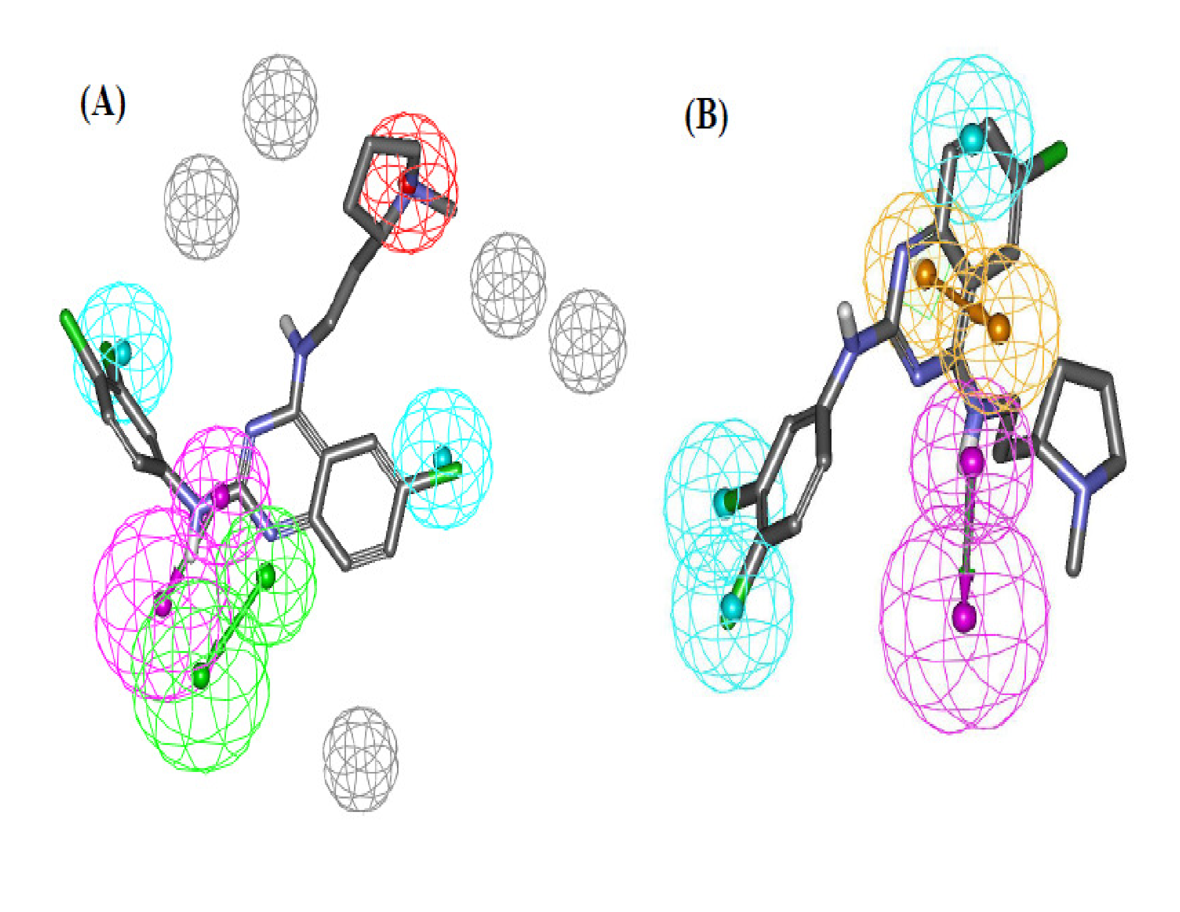
Table 1.10. High-ranking hit molecules with their fit values against Hypo8/31, Hypo9/47, Hypo7/39, their corresponding QSAR estimates from equation (1.10) and their *in vitro* anti-CaMKIIδ bioactivities.

No.	Structure	Fssit values[a]			QSAR Predictions		Experimental	
		Hypo 8/31	Hypo 9/47	Hypo 7/39	log $(1/IC_{50})$	IC_{50} (μM)	% Inhibition at 10 μM	IC_{50} (μM)[b]
89[c,d]	2HCl	10.7	7.3	0	2.31	0.005	100	0.020
90[c,d]	2HCl	6.8	6.5	0	1.62	0.024	100	0.082
91[c,d]	3 HCl	6.81	6.26	0	1.22	0.061	86	0.619
92[c,d]		10.0	8.4	5.4	2.92	0.001	77	2.46
93[d,e]	2HCl	8.8	0	0	-2.47	295	66	1.66

able 1.10. High-ranking hit molecules with their fit values against Hypo8/31, Hypo9/47, Hypo7/39, their orresponding QSAR estimates from equation (1.10) and their *in vitro* anti-CaMKIIδ bioactivities.

No.	Structure	Fit values[a]			QSAR Predictions		Experimental	
		Hypo 8/31	Hypo 9/47	Hypo 7/39	log $(1/IC_{50})$	IC_{50} (μM)	% Inhibition at 10 μM	IC_{50} (μM)[b]
94[c,d]		4.9	6.68	0	-0.24	1.74	56	--
95[e]	3 HCl	5.7	6.2	8.2	1.6	0.025	83.8	1.8
96[e]		10.6	5.64	0	2.66	0.002	80	1.41
97[c,d]		6.48	7.57	0	1.06	0.086	83	1.88
98[f]		0	8.61	0	1.93	0.011	56.43	7.7

[a]Best-fit values calculated by equation (1.4).
[b]Please refer to figure 1.8 and 1.9 for corresponding dose/response curves and correlation coefficients.
[c]Purity ≥ 95% as determined by CHN elemental analysis (see Table 4.3 under Appendix).
[d]Accurate high resolution mass spectrum corresponds to exact calculated mass (see Table 4.3 under Appendix).
[e]Purity < 95% as determined by CHN elemental analysis (see Table 4.3 under Appendix).
[f]No mass spectroscopic or purity information available.



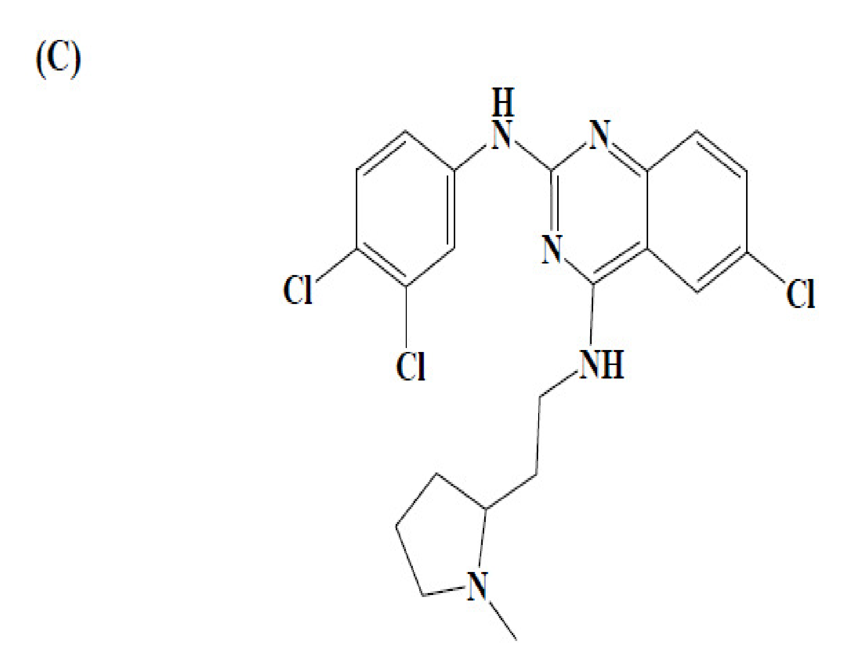


Figure 1.10: Mapping hit compound **89** (IC_{50} = 20 nM, table 1.10) against pharmacophores models **(A)** Hypo8/31; **(B)** Hypo9/47; **(C)** chemical structure of **89** (IC_{50}= 20 nM, table 1.10).

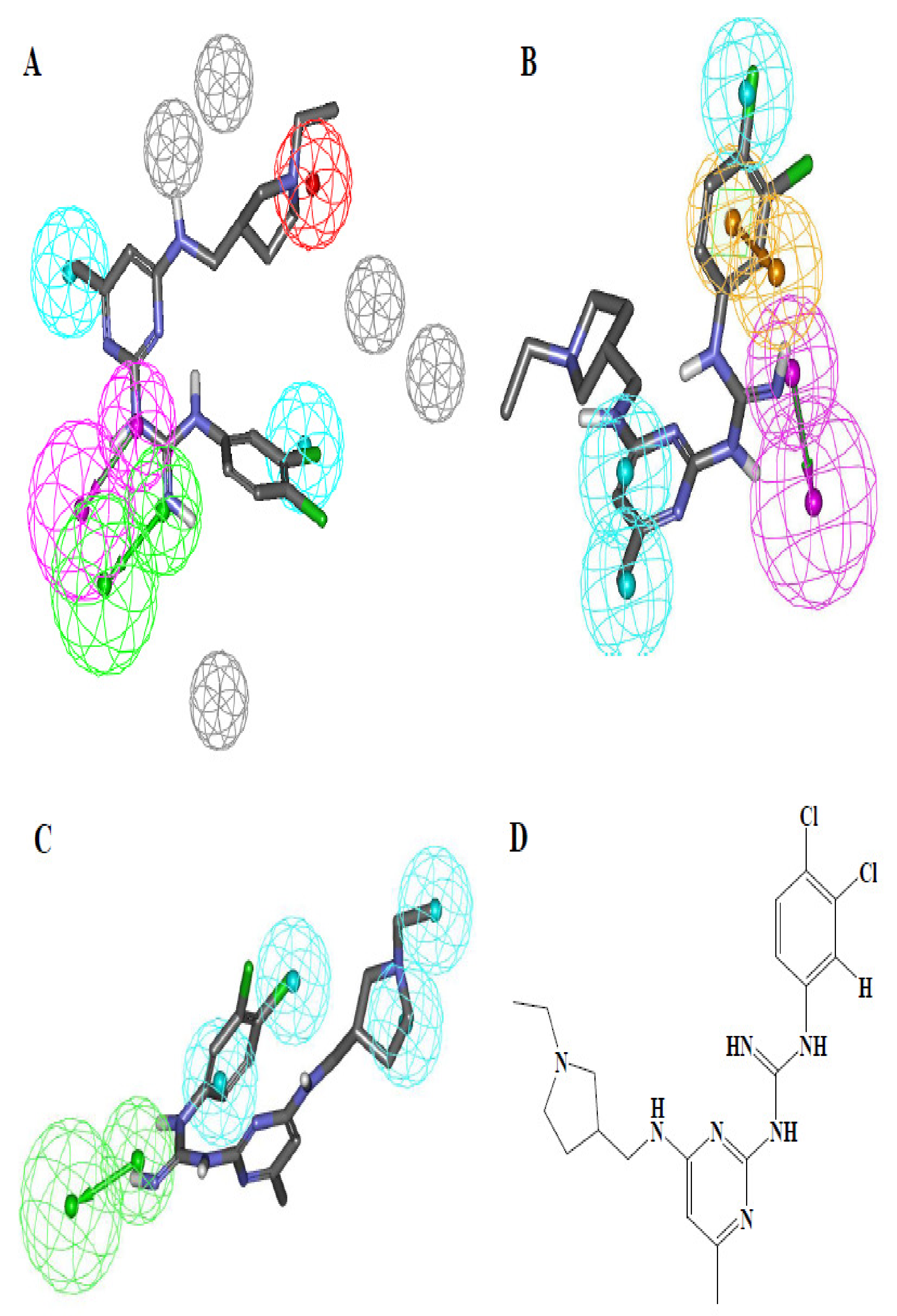


Figure 1.11: Mapping hit compound **92** (IC_{50} = 2.46 µM, table 1.10) against pharmacophores models **(A)** Hypo8/31; **(B)** Hypo9/47; **(C)** Hypo7/39, **(D)** chemical structure of **92** (IC_{50}= 2.46 µM, table 1.10).

It remains to be mentioned that although QSAR predictions were rather accurate with some hit compounds, i.e., they deviated by a maximum of two logarithmic cycles from experimental bioactivities (e.g., **89, 90, 91, 95, and 97**, table 1.10), the experimental IC_{50} values of many other hits differed significantly from QSAR predictions. These errors appear to be because of structural differences between training compounds, used in QSAR and pharmacophore modeling, compared to hit molecules. This discrepancy seems to limit the extrapolatory potential of the QSAR equation. On the other hand, success of QSAR equation in predicting the bioactivities of other hits (e.g. **89** and **90**) is probably because they have close chemical similarity to training molecules (e.g., **63-88**). Furthermore, the fact that we implemented a different bioassay method from that used for training compounds can also explain part of the predicted-to-experimental differences in hit bioactivities. Additionally, our optimal QSAR model has its own weaknesses; in fact an r^2_{PRESS} of 0.600 suggests certain level of uncertainty in predictions.

1.4.6. Pharmacophore-Guided Synthesis of Novel CaMKIIδ inhibitors

The fact that pharmacophore Hypo8/31 was significantly superior to Hypo9/47 and Hypo7/39 vis-à-vis their ROC performances and QSAR slopes, prompted us to employ the former model as template for building novel CaMKIIδ inhibitors. Furthermore, the superb potency of NCI hit **89** (IC_{50} = 20 nM, table 1.10), which exhibits dual central HBA/HBD, two hydrophobic wings and positive ionizable tail, prompted us to envisage triazine-based analogues as potential CaMKIIδ inhibitors. Triazine rings combine two advantages: (i) Facile preparation, and (ii) accessibility to diverse functionalization necessary to satisfy feature-rich nature of Hypo8/31.

Moreover, triazine systems can be easily transformed into amino-triazines securing the appropriate HBA/HBD center required by Hypo8/31. Accordingly, we prepared 14 triazine-based analogues and tested their bioactivities. Table 1.11 shows the prepared compounds, their predicted and experimental bioactivities.

Synthesis proceeded via sequential nucleophilic aromatic replacements of chloro substituents of cyanuric chloride by anilines and amines, as outlined in **Schemes 1.1 – 1.4**. The first step was performed by nucleophilic aromatic substitution attacks by aniline derivatives, as in **scheme 1.1**. The reaction was performed under $^{-}10\ ^{o}C$ - $0\ ^{o}C$ in acetone to give the mono-amino substituted dichlorotriazine derivatives **99-103** in 44-95% yields. Subsequently, the resulting dichlorotriazine derivatives were treated with alkyl amines, namely, pentyl, isopropyl or ethylamines, in THF under ambient conditions to yield the mono-chlorotriazine derivatives **104-113** in 65-99% yields, as in **Scheme 1.2**. Finally, the resulting mono-chlorotriazines were either refluxed with *t*-BOC-protected piperazine in dioxane to produce **114** and **115** (**scheme 1.3**), respectively, or were treated with 1-[2-(dimethylamino)ethyl] piperazine in THF at room temperature to produce **118-127** in 9-64% yields, as in **scheme 1.4**. The *t*-BOC-protected piperazine derivatives **114** and **115** were eventually deprotected by methanolic HCl to yield **116** and **117**, respectively, as shown in **scheme 1.3**. The resulting piperazine-triazine free bases (**118-127**) were dissolved in acetone and eventually precipitated as HCl salts prior to bioassay. This final purification step produced HCl salts at 3 to 5 positions within targeted compounds (**114-127**). These ratios were validated by elemental analyses and in one case, i.e., compound **126**, by argentometric titration of ionic chloride ions (Mohr's method) (Skoog *et al*, 1996).

The synthesized compounds were bioassayed employing CaMKII Assay Kit assay (Cyclex, Japan) (Cyclex CaM kinase II Assay Kit (Cat# CY-1173) Users' Manual,

2009) that was validated by the standard CaMKIIδ inhibitor KN-62 (Davies *et al*, 2000).

Table 1.11 shows the prepared compounds and their corresponding predicted and experimental bioactivities. Clearly from the table, all prepared final products (**114-127**) successfully mapped Hypo8/31 and were predicted by QSAR equation (1.10) to have potent anti-CaMKII bioactivities. However, only 2 illustrated significant experimental potencies (**126 and 127**, table 1.11) ranging from submicromolar to nanomolar IC_{50} values against CaMKII.

Clearly from table 1.11, the experimental bioactivities of prepared compounds differ significantly from their QSAR-based predictions. Nevertheless, we believe it is possible to explain these deviations through the following points: (i) The poor activities of ethyl and isopropyl substituted triazines (e.g., **118**, **119**, **120**, **122**, **123** and **125**) compared to pentyl-substituted analogues (**121**, **126**, and **127**) seem to be because the later can optimally fit Hypo8/31 in at least two distinct poses compared to a single optimal pose for ethyl and isopropyl substituted analogues. Freedom in mapping Hypo8/31 in different poses suggests lesser entropic cost upon binding for the pentyl analogues explaining their higher affinities. Figure 1.12 (A-F) illustrate this point. (ii) The inactivities of *o*-anisidine-based triazines (i.e., **114, 118, 119, 123** and **124**) can be attributed to the formation of intermolecularly hydrogen-bonded 5-membered rings involving the *o*-anisidine OMe and amino-triazine NH. This deprives the NH from its hydrogen-bond donor character leading to the apparent loss of activity of these analogues. Figure 1.12 (**G**, **H**, and **I**) illustrate this point in the case **124**. (iii) The inferior anti-CaMKII bioactivities of unsubstituted piperzine analogues (i.e., **116** and **117**) compared to those dimethyl aminoethyl-substituted piperazine analogues (i.e., **121**, **126** and **127**) can be explained based on the less-than-optimal

mapping of the PosIon feature of the unsubstituted piperazine analogues compared to their *N*-dimethyl aminoethyl-piperazine counterparts. Figure 1.12 clearly explains this point.

99 $R_1 = OCH_3$, $R_2 = CH_3$, R_3=H
100 $R_1 = OCH_3$, R_2 = H, R_3=H
101 $R_1 = CH_3$, R_2 = H, R_3=Cl
102 R_1 = H, R_2 = Cl, R_3=H
103 R_1 = H, $R_2 = NO_2$, R_3=H

Scheme 1.1: Synthesis of 4,6-dichloro-1,3,5-triazin-2-ylamino derivatives. The reactions were conducted in acetone at 0 ^oC.

99 $R_1 = OCH_3$, $R_2 = CH_3$, R_3=H
100 $R_1 = OCH_3$, R_2 = H, R_3=H
101 $R_1 = CH_3$, R_2 = H, R_3=Cl
102 R_1 = H, R_2 = Cl, R_3=H
103 R_1 = H, $R_2 = NO_2$, R_3=H

104 $R_1 = OCH_3$, R_2 = H, R_3=H R_4= Ethyl
105 $R_1 = OCH_3$, R_2 = H, R_3=H, R_4=Isopropyl
106 $R_1 = OCH_3$, $R_2 = CH_3$, R_3=H, R_4=Pentyl
107 $R_1 = OCH_3$, $R_2 = CH_3$, R_3=H, R_4=Isopropyl
108 $R_1 = CH_3$, R_2 = H, R_3=Cl, R_4= Ethyl
109 $R_1 = CH_3$, R_2 = H, R_3=Cl, R_4=Pentyl
110 $R_1 = CH_3$, R_2 = H, R_3=Cl, R_4=Isopropyl
111 R_1 = H, R_2 = Cl, R_3=H, R_4= Ethyl
112 R_1 = H, R_2 = Cl, R_3=H, R_4=Pentyl
113 R_1 = H, $R_2 = NO_2$, R_3=H, R_4=Pentyl

Scheme 1.2: Synthesis of 6-chloro-1,3,5-triazin-2,4-diamino derivatives. The reactions were performed I in THF at 0^oC to room temperature using $(i\text{Pr})_2$NEt as acid scavenger.

104 $R_1 = OCH_3$, $R_2 = H$, R_3=H R_4= Ethyl
109 $R_1 = CH_3$, $R_2 = H$, R_3=Cl, R_4=Pentyl

114 $R_1 = OCH_3$, $R_2 = H$, R_3=H R_4= Ethyl
115 $R_1 = CH_3$, $R_2 = H$, R_3=Cl, R_4=Pentyl

116 $R_1 = OCH_3$, $R_2 = H$, R_3=H R_4= Ethyl
117 R_1 =CH_3 , $R_2 = H$, R_3=Cl, R_4=Pentyl

Scheme 1.3: Synthesis of 4-[4-(phenylamino)-6-alkylamino-[1,3,5]triazin-2-yl]-piperazine via 4-[4-(phenylamino)-6-alkylamino-[1,3,5]triazin-2-yl]-piperazine-1-carboxylic acid tert-butyl ester. **(a)** Reflux with piperazine-1-carboxylic acid tert-butyl ester in dioxane with (*i*Pr)$_2$NEt as acid scavenger, **(b)** methanolic HCl at 0°C to 70°C

104 $R_1 = OCH_3$, $R_2 = H$, R_3=H R_4= Ethyl
105 $R_1 = OCH_3$, $R_2 = H$, R_3=H, R_4=Isopropyl
106 $R_1 = OCH_3$, $R_2 = CH_3$, R_3=H, R_4=Pentyl
107 $R_1 = OCH_3$, $R_2 = CH_3$, R_3=H, R_4=Isopropyl
108 $R_1 = CH_3$, $R_2 = H$, R_3=Cl, R_4= Ethyl
109 $R_1 = CH_3$, $R_2 = H$, R_3=Cl, R_4=Pentyl
110 $R_1 = CH_3$, $R_2 = H$, R_3=Cl, R_4=Isopropyl
111 $R_1 = H$, $R_2 = Cl$, R_3=H, R_4= Ethyl
112 $R_1 = H$, $R_2 = Cl$, R_3=H, R_4=Pentyl
113 $R_1 = H$, $R_2 = NO_2$, R_3=H, R_4=Pentyl

118 $R_1 = OCH_3$, $R_2 = H$, R_3=H R_4= Ethyl
119 $R_1 = OCH_3$, $R_2 = H$, R_3=H, R_4=Isopropyl
120 $R_1 = CH_3$, $R_2 = H$, R_3=Cl, R_4= Ethyl
121 $R_1 = CH_3$, $R_2 = H$, R_3=Cl, R_4=Pentyl
122 $R_1 = CH_3$, $R_2 = H$, R_3=Cl, R_4=Isopropyl
123 $R_1 = OCH_3$, $R_2 = CH_3$, R_3=H, R_4=Isopropyl
124 $R_1 = OCH_3$, $R_2 = CH_3$, R_3=H, R_4=Pentyl
125 $R_1 = H$, $R_2 = Cl$, R_3=H, R_4= Ethyl
126 $R_1 = H$, $R_2 = Cl$, R_3=H, R_4=Pentyl
127 $R_1 = H$, $R_2 = NO_2$, R_3=H, R_4=Pentyl

Scheme 1.4: Synthesis of 6-[4-(2-Dimethylamino-ethyl)-piperazin-1-yl]-N-(3-phenyl)-N'-alkyl-[1,3,5]triazine-2,4-diamine derivatives. **(a)** 1-[2-(dimethylamino) ethyl] piperazine in THF at 25 °C using triethylamine as acid scavenger.

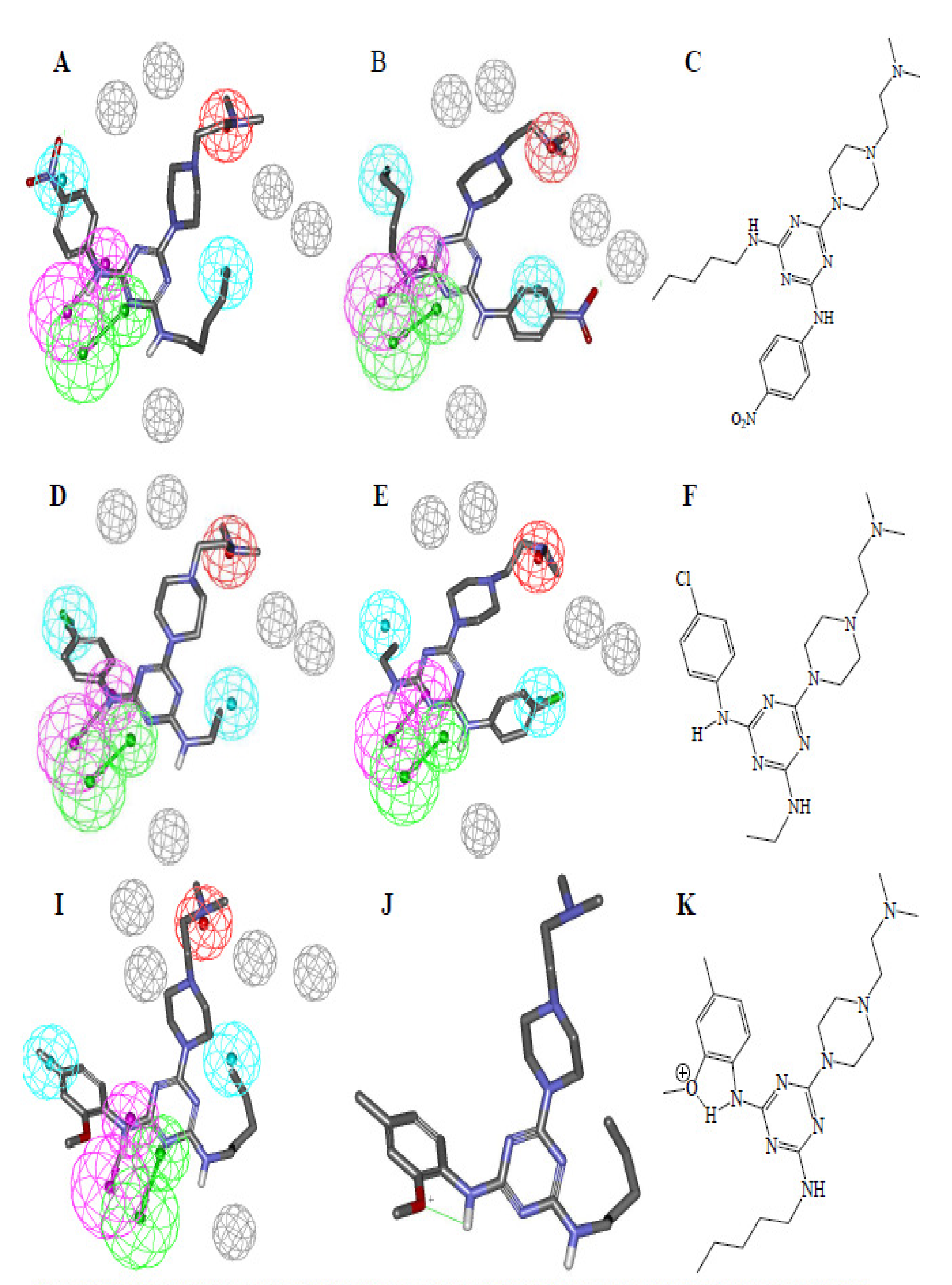


Figure 1.12 (**A**) and (**B**) two optimal alternative mappings of Hypo8/31 against synthesized compound **127** (IC_{50} = 0.154 µM), (**C**) chemical structure of **127**. (**D**) and (**E**) Similar, but suboptimal (the ethyl is not well mapped in **E**), mappings of Hypo8/31 against synthesized compound **125** (inhibition at 10 µM = 6.8%), (**F**) chemical structure of **125**. (**G**) Proposed mapping of **124** ((inhibition at 10 µM = 0%) against Hypo8/31, (**H**) intermolecular hydrogen-bonding involving the *o*-methoxy O and amino-triazine NH in **124** which seems to deprive the amino-triazine NH from fitting the HBD feature in Hypo8/31, (I) chemical structure of **124**.

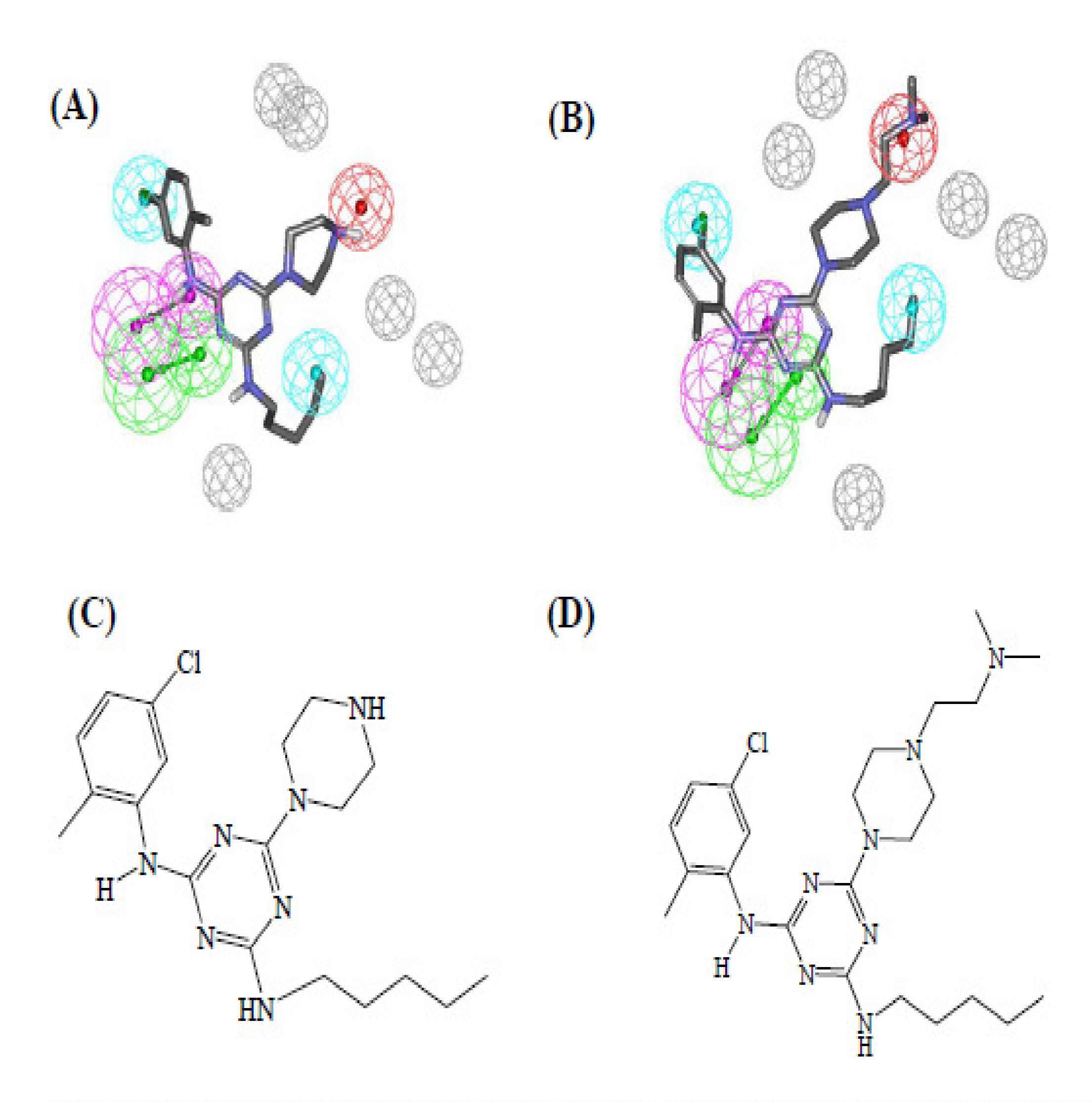


Figure 1.12: **(A)** and **(B)** show Hypo8/31 fitted against synthesized compounds **117** (inhibition at 10 µM = 0%) and **121** (inhibition at 10 µM = 40.16%), respectively, (C) and (D) show the chemical strcutres of **117** and **121**, respectively.

Table 1.11. High-ranking hit molecules with their fit values against Hypo8/31, Hypo9/47, Hypo7/39, their corresponding QSAR estimates from equation (1.10) and their *in vitro* anti-CaMKIIδ bioactivities

No.[a]	Structure	Fit values[b]			QSAR Predictions		Experimental	
		Hypo 8/31	Hypo 9/47	Hypo 7/39	log (1/IC_{50})	IC_{50} (μM)	% Inhibition at 10 μM	IC_{50} (μM)
116		11.64	4.07	0	3.90	0.0001	0	--
117		8.15	5.46	0	3.26	0.0005	0	--
118		8.84	7.07	0	3.17	0.0007	0	--
119		7.44	7.11	0	2.87	0.0014	0	--

Table 1.11. High-ranking hit molecules with their fit values against Hypo8/31, Hypo9/47, Hypo7/39, their corresponding QSAR estimates from equation (1.10) and their *in vitro* anti-CaMKIIδ bioactivities

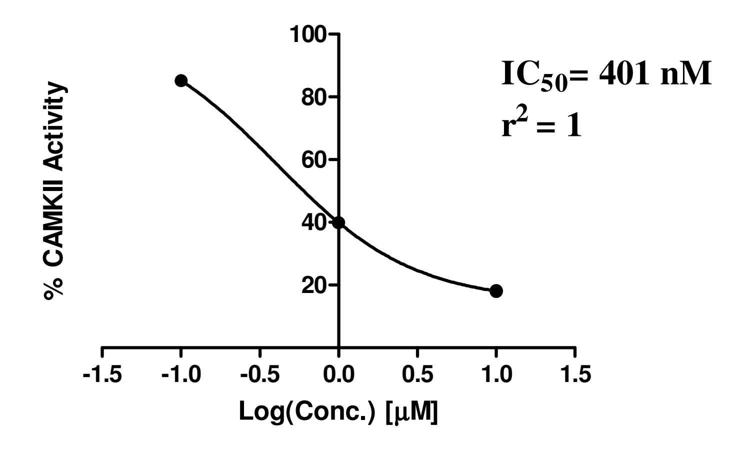
No.[a]	Structure	Fit values[b]			QSAR Predictions		Experimental	
		Hypo 8/31	Hypo 9/47	Hypo 7/39	log $(1/IC_{50})$	IC_{50} (μM)	% Inhibition at 10 μM	IC_{50} (μM)
120		9.25	5.74	0	2.89	0.0013	0	--
121		10.72	7.62	0	3.33	0.0005	40.16	--
122		8.2	5.63	0	2.65	0.0023	1.35	--
123		7.48	6.82	0	2.87	0.0013	0	--
124		10.74	7.82	4.35	3.95	0.0001	0	--

Table 1.11. High-ranking hit molecules with their fit values against Hypo8/31, Hypo9/47, Hypo7/39, their corresponding QSAR estimates from equation (1.10) and their *in vitro* anti-CaMKIIδ bioactivities

No.[a]	Structure	Fit values[b]			QSAR Predictions		Experimental	
		Hypo 8/31	Hypo 9/47	Hypo 7/39	log $(1/IC_{50})$	IC_{50} (µM)	% Inhibition at 10 µM	IC_{50} (µM)
125		7.74	6.40	0	3.16	0.0007	6.88	--
126		10.18	8.05	0	3.58	0.0003	82	0.401[c]
127		7.04	7.03	0	1.57	0.0271	82	0.154[c]
KN-62				-----------			85.5	0.380[d]

[a]See experimental section 1.3.2.1 for detailed characterization (NMR, IR, HRMS, Elemental analysis)
[b]Best-fit values calculated by equation (1.4)
[c]Please refer to figure 1.13 for corresponding dose/response curves and correlation coefficients.
[d]Please refer to figure 1.9 for corresponding dose/response curve and correlation coefficient.

(A)



(B)

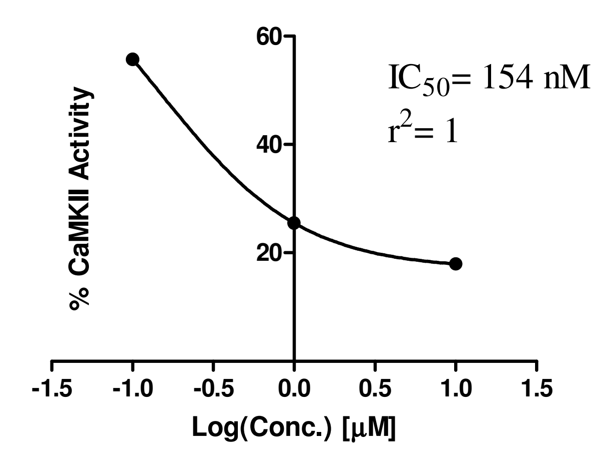


Figure 1.13: Dose-response plots of highest active synthesized compounds (A) **126**, and (B) **127**. Data fitting was performed using GraphPad Prism (version 5.04) via fitting against sigmoidal dose-inhibition model.

1.5 Conclusion

CaMKIIδ inhibitors are currently considered as potential treatments for cardiovascular disease. The pharmacophoric space of CaMKIIδ inhibitors was explored via nine diverse sets of inhibitors and using CATALYST-HYPOGEN to identify high quality binding model(s). Subsequently, genetic algorithm and multiple linear regression analysis were employed to access optimal QSAR model capable of explaining anti-CaMKIIδ bioactivity variation across 88 collected CaMKIIδ inhibitors. Three orthogonal pharmacophoric models emerged in the QSAR equation suggesting the existence of at least three distinct binding modes accessible to ligands within CaMKIIδ binding pocket. The QSAR equation and the associated pharmacophoric models were experimentally validated by the identification of several CaMKIIδ inhibitors retrieved via *in silico* screening. The QSAR equation and the associated pharmacophoric models were used to guide synthetic exploration of a triazine based new series of CaMKIIδ inhibitors that resulted in several novel nanomolar CaMKIIδ inhibitors.

.

Chapter Two

Elaborate Ligand-Based Modeling Reveal New Submicromolar Rho Kinase Inhibitors

2.1 Chapter Abstract

Rho Kinase (ROCKII) has been recently implicated in several cardiovascular diseases prompting several attempts to discover and optimize new ROCKII inhibitors. Towards this end we explored the pharmacophoric space of 138 ROCKII inhibitors using six diverse sets of inhibitors to identify high quality pharmacophores. The pharmacophoric models were subsequently allowed to compete with each other with the context of QSAR modeling. Towards this end, genetic algorithm and multiple linear regression analysis were employed to select an optimal combination of pharmacophoric models and 2D physicochemical descriptors capable of accessing self-consistent quantitative structure-activity relationship (QSAR) of optimal predictive potential (r_{111} = 0.82, F = 18.752, r^2_{LOO} = 0.679, r^2_{PRESS} against 27 external test inhibitors = 0.535). Two orthogonal pharmacophores emerged in the QSAR equation suggesting the existence of at least two binding modes accessible to ligands within ROCKII binding pocket. Receiver operating characteristic (ROC) curve analyses established the validity of QSAR-selected pharmacophores. We employed the pharmacophoric models and associated QSAR equation to screen the national cancer institute (NCI) list of compounds Eight submicromolar ROCKII inhibitors were identified. The most potent gave IC_{50} values of 0.7 and 1 uM.

2.2. Introduction

Rho-kinases (or Rho associated protein kinases, ROCKs) are downstream targets of the Rho G-proteins. They are involved in the regulation of cytoskeletal reorganization and gene expression, as in Figure 2.1 (Shimokawa *et al*, 2007).

There are two isoforms of ROCK: ROCK I (or ROKβ) and ROCK II (or ROKα) (Olson, 2008). Both isoforms are highly homologous, sharing 65% homology in amino acid sequence and 92% homology in their kinase domains. However, although both isoforms are ubiquitously expressed, ROCK II is highly expressed in the brain and the heart, whereas ROCK I is expressed preferentially in the lung, liver, spleen, kidney and testis. Nevertheless, evidence of the functional differences between ROCK I and ROCK II is still lacking (Kumar *et al*, 2007)..

The catalytic kinase domain of the two ROCK isoforms is located at the amino terminus, followed by coiled-coil-forming region that encompasses the Rho-binding domain (RBD), which is followed in turn by pleckstrin-homology domain (PH). The cysteine-rich repeat domain (CRD) is present at the carboxyl terminus (Muller *et al,* 2005).

The carboxyl terminus of ROCKs folds back onto the kinase domain, thereby forming an auto-inhibitory loop that maintains the enzyme in an inactive state. Binding of active Rho to ROCKs disrupts this negative regulatory interaction resulting in activation of the enzyme in response to extracellular signals (Muller *et al,* 2005).

ROCK II is involved in the regulation of vascular tone, endothelial function, inflammation and remodeling of cardiac cells (Muller *et al,* 2005, Offermanns &

Wettschurech, 2002). In particular, abnormal activation of the Rho/ROCK II pathway has been shown to play a role in hypertension (Offermanns & Wettschurech, 2002).

ROCK II appears to increase the force and velocity of actinomyosin cross-bridging in smooth muscle and nonmuscle cells by maintaining the phosphorylated state of myosin light chain (MLC). This is achieved through direct phosphorylation of MLC and via inhibiting myosin light chain phosphatase (MLCPh) (Kumar *et al*, 2007), see **Figure 2.1**.

Accordingly, inhibition of ROCK II should have beneficial effects in a variety of cardiovascular disorders including hypertension, atherosclerosis, ischemia–reperfusion injury, stroke, myocardial hypertrophy, heart failure, cardiac allograft vasculopathy and vein graft disease (Dong *et al*, 2010). The therapeutic significance of ROCK II in cardiovascular medicine has recently attracted a great deal of attention (Shimokawa *et al*, 2007).

The main focus of recent efforts towards the development of new ROCK II inhibitors concentrate on structure-based ligand design (Takami *et al*, 2004, Iwakubo *et al*, 2007 , a&b , Ho *et al*, 2010). To date, several human ROCK II complexes have been resolved by X-ray crystallography and solution NMR, e.g., PDB codes: 2H9V, 2ROW, 2ROV, 2F2U.(Yamaguchi *et al*, 2006 , a&b)

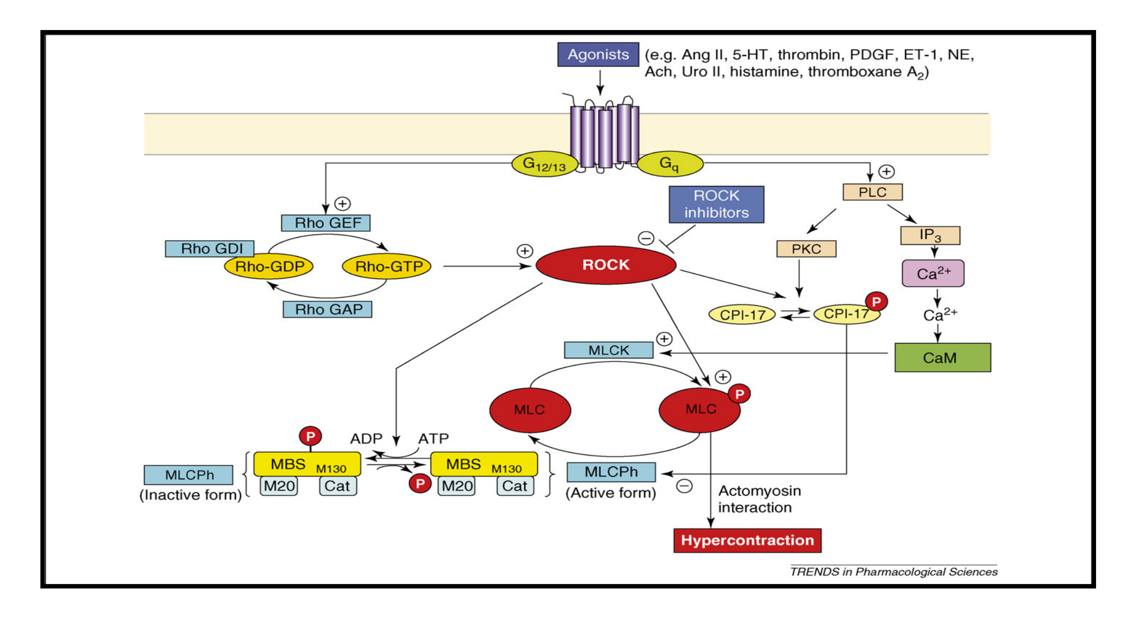


Figure 2.1: The Rho–ROCK signaling pathway in vascular smooth muscle cell contraction (Shimokawa *et al,* 2007). Abbreviations: Ach, acetylcholine; Ang II, angiotensin II; Cat, catalytic subunit; ET-1, endothelin-1; IP3, inositol (1,4,5)-trisphosphate; M20, 20-kDa subunit; NE, norepinephrine; PLC, phospholipase C; PDGF, platelet-derived growth factor; Uro II, urotensin II. Stimulation is denoted by +; inhibition is denoted by -.

We employed the CATALYST-HYPOGEN module of the software package Discovery Studio (Discovery Studio 2.5.5 User Guide., 2010) to construct plausible binding hypotheses for a diverse list of ROCKII inhibitors (Takami *et al*, 2004, Iwakubo *et al*, 2007 , a&b).Subsequently, genetic function algorithm (GFA) and multiple linear regression (MLR) analyses were employed to search for an optimal QSAR that combine high-quality binding pharmacophores with other molecular descriptors and capable of explaining bioactivity variation across a collection of diverse ROCK II inhibitors. The

optimal pharmacophores were subsequently used as 3D search queries to screen the national cancer institute (NCI) list of compounds for new ROCK II inhibitory leads.

CATALYST-HYPOGEN models drug-receptor interaction using information derived only from the ligand structure. It identifies a 3D array of a maximum of five chemical features common to active training molecules, which provides a relative alignment for each input molecule consistent with their binding to a proposed common receptor site. The chemical features considered can be hydrogen bond donors and acceptors (HBDs and HBAs), aliphatic and aromatic hydrophobes (Hbic), positive and negative ionizable (PosIon and NegIon) groups and aromatic planes (RingArom). The conformational flexibility of training ligands is modeled by creating multiple conformers, judiciously prepared to emphasize representative coverage over a specified energy range. CATALYST pharmacophores have been used as 3D queries for database searching and in 3D-QSAR studies (Taha *et al*, 2008 , a, b & c, Taha *et al,* 2007 , Al-masri *et al,* 2008 , Al-Nadaf *et al*, 2010 , Abu-Hammad *et al,* 2009 , Abu Khalaf *et al*, 2010 , Al-Sha'er and Taha, 2010 , a&b, Taha *et al,* 2010, (Abu Khalaf *et al*, DOI 10.1007/s00894-010-0737-1 , Abdula *et al*, 2011 , Van Drie , 2003 , Poptodorov *et al,* 2006).

2.3. Experimental

2.3.1. Molecular Modeling

2.3.1.1. Software and Hardware

The following software packages were utilized in the present research.

- CATALYST (Version 4.11), Accelrys Inc. (www.accelrys.com), USA.
- CERIUS2 (Version 4.10), Accelrys Inc. (www.accelrys.com), USA.
- CS ChemDraw Ultra 6.0, Cambridge Soft Corp. (http:// www.cambridgesoft.Com), USA.
- Discovery Studio 2.5, Accelrys Inc. (www.accelrys.com), USA.

Pharmacophore and QSAR modeling studies were performed using CATALYST (HYPOGEN module), CERIUS2 software suites from Accelrys Inc. (San Diego, California, www.accelrys.com) and Discovery Studio. Structure drawing was performed employing ChemDraw Ultra 6.0 which was installed on a Pentium 4 PC.

2.3.1.2. Data Set in ROCK II

The structures of 138 ROCK II inhibitors (Table 2.1) were collected from recently published literature (Takami *et al*, 2004, Iwakubo *et al*, 2007, a&b). Although the collected inhibitors were gathered from three separate articles, they were bioassayed employing the same bioassay methodology. The bioactivities were expressed as the concentrations of the test compounds that inhibited the activity of ROCK II by 50% (IC_{50} in μM). The logarithm of measured IC_{50} values were used in QSAR and pharmacophores analyses, thus correlating the data linear to the free energy change.

In cases where IC_{50} is expressed as being higher than 10 μM (e.g., compounds **4**, **5**, **24**, **44** and **58**, see **table 2.1**) it was assumed it equals 35 μM. This assumption is necessary to allow 3.5 log cycles separation between the most potent training compounds (i.e., **113** and **120**) and the least active training compounds (i.e., **4**, **5**, **24**, **44** and **58**). This is necessary for CATALYST pharmacophores modeling. Moreover, this assumption is necessary to allow statistical correlation and QSAR analysis since regression requires exact values. The logarithmic transformation of IC_{50} values should minimize any potential errors resulting from this assumption.

The two-dimensional (2D) chemical structures of the inhibitors were sketched using ChemDraw Ultra and saved in MDL-molfile format. Subsequently, they were imported into CATALYST, converted into corresponding standard 3D structures and energy minimized to the closest local minimum using the molecular mechanics CHARMm force field implemented in CATALYST. The resulting 3D structures were utilized as starting conformers for CATALYST conformational analysis.

Table 2.1: The structures of ROCK II inhibitors utilized in modeling

No.	Reference [a]	R_1	R_2	R_3	R_4	R_5	R_6	R_7	X	Y	IC_{50} (μM)
1	1	H	H	H	H	Cl	Cl	Cl	NH	NH	0.2
2	1	H	H	H	H	Cl	Cl	H	NH	NH	0.9
3	1	H	H	H	H	Cl	H	Cl	NH	NH	>10
4	1	H	H	H	H	Cl	H	H	NH	NH	>10
5	1	H	H	H	H	H	H	H	NH	NH	>10
6	1	H	H	H	H	F	F	H	NH	NH	2
7	1	H	H	H	H	CF_3	F	H	NH	NH	13.6
8	1	H	H	H	H	Cl	Cl	H	NH	$NHCH_2$	0.8
9	1	H	H	H	H	Cl	H	Cl	NH	$NHCH_2$	4.1
10	1	H	H	H	H	Cl	H	H	NH	$NHCH_2$	7.2
11	1	H	H	H	H	H	H	H	NH	$NHCH_2$	10.2
12	1	H	H	H	H	F	F	H	NH	$NHCH_2$	7
13[b]	1	H	H	H	H	CF_3	F	H	NH	$NHCH_2$	2.2
14	1	H	H	H	H	Cl	F	Cl	NH	CH_2O	0.5
15	1	H	H	H	H	Cl	H	Cl	NH	CH_2O	0.9
16	1	H	H	H	NH_2	Cl	H	Cl	NH	NH	>10
17	1	H	H	Cl	Cl	Cl	H	Cl	NH	NH	>10
18	1	Cl	Cl	H	H	Cl	H	Cl	NH	NH	>10
19	1	Cl	H	H	H	Cl	H	Cl	NH	NH	>10
20[b]	1	NH_2	H	H	H	Cl	H	Cl	NH	NH	>10
21	1	H	H	H	H	Cl	H	Cl	NCH_3	NCH_3	>10
22	1	H	H	H	H	H	H	Cl	CH_2O	$HNCH_2$	>10
23	1	H	H	NH	H	Cl	H	Cl	NH	$HNCH_2$	>10
24	1	Cl	H	H	H	Cl	H	Cl	NH	CH_2O	>10

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	R_4	R_5	X	IC_{50} (μM)
25	1	H	H	H	H		N	0.5
26[b]	1	H	H	H	H		N	0.9
27[b]	1	H	H	H	H		N	0.5
28	1	H	H	H	H		N	1
29	1	H	H	H	H		N	0.3
30	1	H	H	H	H		N	0.3
31	1	H	H	H	H		N	0.7
32	1	H	H	H	H		N	0.2

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference [a]	R_1	R_2	R_3	R_4	R_5	X	IC_{50} (µM)
33	1	H	H	H	H		CH	>10
34	1	F	H	H	H		N	>10
35	1	OMe	H	H	H		N	>10
36	1	H	COMe	H	H		N	>10
37[b]	1	H	Et	H	H		N	>10
38	1	H	H	Me	H		N	>10
39	1	H	H	H	Me		N	>10

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	R_4	R_5	X	IC_{50} (µM)
40	1	H	H	H	H	N	N	0.8
41	1	H	H	H	H	N	N	0.02
42	1	H	H	H	H	NH	N	0.1
43	1	H	H	H	H	N	N	0.03
44[b]	1	H	H	H	Me	N	N	>10

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	R_4	R_5	X	IC_{50} (µM)
45[b]	1	H	H	H	H		N	>10
46	1	H	H	H	H		N	>10
47	1	H	H	H	H		N	>10
48	1	H	H	H	H		N	>10
49[b]	1	H	H	H	H		N	>10

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference [a]	R_1	R_2	X	Y	IC_{50} (µM)
50[b]	1	*	H	N	CH	0.1
51[b]	1	*	H	N	CH	0.3
52[b]	1	*–$NHCH_3$	H	CH	CH	0.2
53	1	H	*	CH	N	0.2
54	1	H	*	CH	N	0.1
55	3	*	H	N	CH	0.21
56	3	H	*	CH	N	0.27

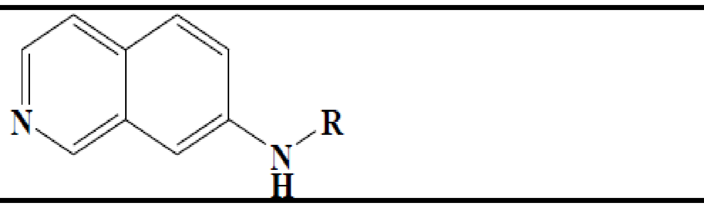
No.	Reference [a]	R_1	R_2	X	Y	IC_{50} (µM)
57	1	*	H	N	CH	>10
58	1	H	*	CH	N	>10

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	n	IC_{50} (μM
59	1	--	COMe	--	0	>10
60	1	Cl	--	Cl	0	>10
61	1	Cl	Cl	Cl	0	>10
62	1	F	--	F	1	>10
63[b]	1	Cl	--	Cl	1	>10

No.	Reference[a]	R	IC_{50} (μM)
64	3		0.18
65	3		0.18
66	3		0.195
67[b]	3		0.07

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.



No.	Reference[a]	R	IC_{50} (μM)
68	3		0.44
69	3		0.58

No.	Reference[a]	R_1	R_2	X	Y	IC_{50} (μM)
70[b]	1	---	H	NH	CH_2	0.4
71	1	2-Cl	H	NH	CH_2	0.3
72	1	2,6-Cl	H	NH	CH_2	0.1
73	1	2,6-Cl	Cl	NH	CH_2	0.1
74	1	2,6-Cl	H	O	CH_2	>10
75[b]	1	2,6-Cl	H	CH_2	O	>10

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R	IC_{50} (μM)
76[b]	1		8.3
77	1		0.9
78	1	Cl, Cl	1.9
79	1	N	>10
80	1	N, N	>10

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

R2 — Linker — R1

R3

No.	Reference[a]	R_1	R_2	R_3	Linker	IC_{50} (μM)
81	2	Cl, *, Cl	H	H	*, H N, H N, O, *	0.02
82	2	F, *, F	H	H	*, H N, H N, O, *	0.085
83[b]	2	Cl, Cl, *	H	H	*, H N, H N, O, *	0.22
84	2	*	H	H	*, H N, N, *	7.9
85	2	*, F	H	H	*—NH, N, *	>10
86	2	*	H	H	O, *—NH, N, *	0.136
87	2	*	H	H	HN, *, O, N, *	0.54

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Linker	IC_{50} (μM)
88	2		H	H		0.018
89	2		H	H		>10
90	2		H	H		0.1
91[b]	2		H	H		0.01
92[b]	2		H	H		>10
93[b]	2		H	CH3		0.15
94	2	Cl	H	H		0.44

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Linker	IC_{50} (µM)
95	2	Cl	H	H	HN * N *	0.29
96	2	* Cl	H	H	HN * N *	0.29
97	2	Cl *	H	H	H N * N *	0.045
98[b]	2	Cl *	H	H	H N * N *	0.02
99	2	* Cl	H	H	H N * N *	0.024
100[b]	2	Cl *	H	H	H * N N *	0.013

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

R2, N, N, R3, Linker, R1

No.	Reference [a]	R_1	R_2	R_3	Linker	IC_{50} (μM)
101	2	Cl	H	H	H N N	0.019
102	2	Cl	H	H	H N N	0.011
103[b]	2	CH_3	H	H	H N N	0.780
104	2		H	H	H N N	0.34
105	2	F	H	H	H N N	0.025
106	2	F F	H	H	H N N	0.01
107	2		H	H	H N N	0.009

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

R_2, N, N, R_3, [Linker]–R_1

No.	Reference[a]	R_1	R_2	R_3	Linker	IC_{50} (μM)
108	2	* O	H	H	* H N N *	0.016
109	2	* O O	H	H	* H N N *	0.08
110[b]	2	* H_2N	H	H	* H N N *	0.085
111[b]	2	* NH_2	H	H	* H N N *	0.035
112	2	* NH_2	H	H	* H N N *	0.01
113	2	* O N^+ O^-	H	H	* H N N *	0.003

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Linker	IC_{50} (μM)
114	2		H	H		0.087
115	2		H	H		0.24
116	2		H	H		0.072
117	2		H	H		0.19
118	2		H	H		0.011
119	2		H	H		0.005

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No.	Reference[a]	R_1	R_2	R_3	Linker	IC_{50} (μM)
120	2		H	H		0.003
121	2		H	H		0.065
122	2		H	H		0.18
123	2		H	H		0.32
124	2		H	H		0.52

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No	Reference[a]	X	n	IC_{50} (μM)
125	3	2-Cl	1	0.21
126	3	3-Cl	1	0.265
127[b]	3	4-Cl	1	0.185
128	3	2-Cl	0	0.1
129	3	3-Cl	0	0.065
130	3	4-Cl	0	0.065
131	3	4-F	0	0.095
132	3	2,6-F	0	0.075
133	3	2-NH2	0	0.205
134[b]	3	3-NH2	0	0.275
135	3	4-NH2	0	0.235

Table 2.1: The structures of ROCK II inhibitors utilized in modeling.

No	Reference[a]		IC_{50} (µM)
136 (Fasudil)	3		0.205
137	1		>10
138	1		9

[a]References: 1: Takami *et al,* 2004; 2: Iwakubo *et al,* 2007, a , 3: Iwakubo *et al,* 2007
[b]These compounds were employed as external test subset in QSAR modeling
* Connection point

2.3.1.3 Conformational Analysis

The molecular flexibilities of the collected compounds were taken into account by considering each compound as a collection of conformers representing different areas of the conformational space accessible to the molecule within a given energy range. Accordingly, the conformational space of each inhibitor (1-138, Table 2.1) was explored adopting the “best conformer generation” option within CATALYST (CATALYST 4.11 users’ manual, 2005). based on the generalized CHARMm force field implemented in the program. Default parameters were employed in the conformation generation procedure of training compounds, i.e., a conformational ensemble was generated with an energy threshold of 20 kcal/mol from the local minimized structure at which has the lowest energy level and a maximum limit of 250 conformers per molecule (CATALYST 4.11 users’ manual, 2005).

2.3.1.4 Pharmacophoric Hypotheses Generation

All 138 molecules with their associated conformational models were regrouped into a spreadsheet. The biological data of the inhibitors were reported with an “ ainty” value of 3, which means that the actual bioactivity of a particular inhibitor is assumed to be situated somewhere in an interval ranging from one-third to three-times the reported bioactivity value of that inhibitor (Kurogi and Güner, 2001, Sutter *et al*, 2000). Subsequently, six structurally diverse training subsets: sets I, II, III, IV, V and VI (Table 2.2) were carefully selected from the collection for pharmacophore modeling. Typically, CATALYST requires informative training sets that include at least 16

compounds of evenly spread bioactivities over at least three and a half logarithmic cycles. Lesser training lists could lead to chance correlation and thus faulty models.

Table 2.2: Training subsets employed in exploring the pharmacophoric space of ROCK II inhibitors, numbers correspond to compounds in Table 2.1..

subsets	Most active[a]	Moderate active	Least active[b]	No. of compounds
I	**52, 57, 65, 82, 87, 91, 108, 121, 130, 132, 133**	**6, 14, 26, 27, 30, 51, 64, 69**	**45**	22
II	**41, 108, 109, 113, 114, 121**	**27, 29, 30, 52, 64, 69, 83, 91, 124, 130, 132**	**4, 5, 7, 11, 56, 58, 59, 62**	25
III	**89, 92, 99, 103, 108, 120, 121**	**25, 26, 30, 50, 54, 63, 65, 87, 91, 128, 131**	**5, 45, 49, 62**	22
IV	**32, 50, 53, 54, 63, 65, 87, 111, 119, 129, 134, 136**	**64, 30, 31, 51, 55, 135**	**5, 11, 45, 46**	22
V	**99, 103, 107, 108, 121**	**25, 26, 29, 30, 42, 43, 50, 54, 63, 65, 130**	**3, 4, 5, 19, 22, 62**	22
VI	**32, 50, 63, 94, 107, 109, 114, 115, 118, 132, 133, 136**	**34, 64**	**11, 42, 45, 49, 62**	19

[a]Potency categories as defined by Equation (2.1). [b]Potency categories as defined by Equation (2.2)

The selected training sets were utilized to conduct 36 modeling runs to explore the pharmacophoric space of ROCK II inhibitors (Table 2.3). The exploration process included altering interfeature spacing parameter (100 and 300 picometers) and the maximum number of allowed features in the resulting pharmacophore hypotheses, i.e., they were allowed to vary from 4 to 5 for first and second runs and from 5 to 5 for third and fourth runs of each training set (Table 2.3).

Pharmacophore modeling employing CATALYST proceeds through three successive phases: the constructive phase, subtractive phase and optimization phase. During the constructive phase, CATALYST generates common conformational alignments among the most-active training compounds. Only molecular alignments based on a maximum of five chemical features are considered. The program identifies a particular compound as being within the most active category if it satisfies equation (2.1) (Poptodorov *et al*., 2006; Kurogi and Güner, 2001; Li *et al*., 2000, Sutter *et al,* 2000, Bersuker *et al*., 2000).

$$(\text{MAct x UncMAct}) - (\text{Act / UncAct}) > 0.0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.1)$$

Where "MAct" is the activity of the most active compound in the training set, "Unc" is the uncertainty of the compounds and "Act" is the activity of the training compounds under question. However, if there are more than eight most-active inhibitors, only the top eight are used.

In the subsequent subtractive phase, CATALYST eliminates some hypotheses that fit inactive training compounds. A particular training compound is defined as being inactive if it satisfies equation (2.2): (Poptodorov *et al*., 2006; Kurogi and Güner, 2001; Li *et al*., 2000, Sutter *et al,* 2000, Bersuker *et al*., 2000).

$$\text{Log (Act)} - \log (\text{MAct}) > 3.5 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.2)$$

However, in the optimization phase, CATALYST applies fine perturbations in the form of vectored feature rotation, adding new feature and/or removing a feature, to selected hypotheses that survived the subtractive phase, in an attempt to find new models of enhanced bioactivity/mapping correlation, i.e., improved 3D-QSAR properties (Poptodorov *et al*., 2006; Kurogi and Güner, 2001; Li *et al*., 2000, Sutter *et al,* 2000, Bersuker *et al*., 2000).

Table 2.2 shows the most active, least active and intermediately active compounds of each training subset as categorized by equations (2.1) and (2.2). Eventually, CATALYST selects the highest-ranking models (10 by default) and presents them as the optimal pharmacophore hypotheses resulting from the particular automatic modeling run. Our pharmacophore exploration efforts (36 automatic runs, Tables 2.2 and 2.3) culminated in 336 pharmacophore models of variable qualities.

Table 2.3: Training sets and CATALYST run parameters employed for exploring ROCKII pharmacophoric space

Run Number	Training set[a]	Number of compounds	Selected features[b]	Max-Min Features[c]	Spacing[d]
1	I	22	Hbic (0-3), HBA (0-3), HBDN (0-3),	4-5	100
2		22	Hbic (0-3), HBA (0-3), HBDN (0-3),	4-5	300
3		22	Hbic (0-3), HBA (0-3), HBDN (0-3),	5-5	100
4		22	Hbic (0-3), HBA (0-3), HBDN (0-3),	5-5	300
5		22	Hbic (0-3), HBA (0-3), HBDN (2-3),	4-5	100
6		22	Hbic (0-3), HBA (0-3), HBDN (2-3),	4-5	300
7		22	Hbic (0-3), HBA (0-3), HBDN (2-3),	5-5	100
8		22	Hbic (0-3), HBA (0-3), HBDN (2-3),	5-5	300
9	II	25	Hbic (0-3), HBA (0-3), HBD (0-3),	4-5	100
10		25	Hbic (0-3), HBA (0-3), HBD (0-3),	4-5	300
11		25	Hbic (0-3), HBA (0-3), HBD (0-3),	5-5	100
12		25	Hbic (0-3), HBA (0-3), HBD (0-3),	5-5	300
13	III	22	Hbic (0-3), HBA (0-3), HBD (0-3),	4-5	100
14		22	Hbic (0-3), HBA (0-3), HBD (0-3),	4-5	300
15		22	Hbic (0-3), HBA (0-3), HBD (0-3),	5-5	100
16		22	Hbic (0-3), HBA (0-3), HBD (0-3),	5-5	300
17	IV	22	Hbic (0-3), HBA (0-3), HBD (0-3),	4-5	100
18		22	Hbic (0-3), HBA (0-3), HBD (0-3),	4-5	300
19		22	Hbic (0-3), HBA (0-3), HBD (0-3),	5-5	100
20		22	Hbic (0-3), HBA (0-3), HBD (0-3),	5-5	300
21		22	Hbic (0-3), HBA (0-3), HBD (2-3),	4-5	100
22		22	Hbic (0-3), HBA (0-3), HBD (2-3),	4-5	300
23		22	Hbic (0-3), HBA (0-3), HBD (2-3),	5-5	100
24		22	Hbic (0-3), HBA (0-3), HBD (2-3),	5-5	300
25	V	22	Hbic (0-3), HBA (0-3), HBD (0-3),	4-5	100
26		22	Hbic (0-3), HBA (0-3), HBD (0-3),	4-5	300
27		22	Hbic (0-3), HBA (0-3), HBD (0-3),	5-5	100
28		22	Hbic (0-3), HBA (0-3), HBD (0-3),	5-5	300
29	VI	19	Hbic (0-3), HBAF (0-3), HBDN (0-3),	4-5	100
30		19	Hbic (0-3), HBAF (0-3), HBDN (0-3),	4-5	300
31		19	Hbic (0-3), HBAF (0-3), HBDN (0-3),	5-5	100
32		19	Hbic (0-3), HBAF (0-3), HBDN (0-3),	5-5	300
33		19	HBAF (1-1), HBDN (1-1), RingArom (1-	4-4	100
34		19	HBAF (1-1), HBDN (1-1), RingArom (1-	4-4	100[e]
35		19	Hbic (0-1),HBAF (1-1), HbicArom (1-1),	4-4	100
36		19	Hbic (0-1),HBAF (1-1), HbicArom (1-1),	4-4	100[e]

[a]Training subsets as in table B under Supplementary Materials.
[b]HBA: Hydrogen Bond Acceptor, HBD: Hydrogen Bond Donor, HBDN: Hydrogen Bond Donors including ionizable nitrogen atoms, HBAF: Hydrogen Bond Acceptors including fluorine atoms, RingArom: Ring Aromatic, Hbic: Hydrophobic, HbicArom : Hydrophobic Aromatic, PosIon: PosIonizable. Numbers in brackets refer to the allowed ranges of corresponding features in each pharmacophore modeling run.
[c]Max-Min refers to the allowed range of pharmacophoric features in each model.
[d]Spacing refers to the maximum interfeature distance in picometers. Other parameters

2.3.1.5 Assessment of the Generated Hypotheses

When generating hypotheses, CATALYST attempts to minimize a cost function consisting of three terms: Weight cost, Error cost and Configuration cost (CATALYST 4.11 users' manual, 2005; Poptodorov *et al*., 2006; Kurogi and Güner, 2001; Li *et al*., 2000; Sutter *et al,* 2000; Bersuker *et al,* 2000). Weight cost is a value that increases as the feature weight in a model deviates from an ideal value of 2. The deviation between the estimated activities of the training set and their experimentally determined values adds to the error cost. The activity of any compound can be estimated from a particular hypothesis through equation (2.3) (CATALYST 4.11 users' manual, 2005).

$$\text{Log (Estimated Activity)} = \text{I} + \text{Fit} \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\; (2.3)$$

Where, I = the intercept of the regression line obtained by plotting the log of the biological activity of the training set compounds against the Fit values of the training compounds. The Fit value for any compound is obtained automatically employing equation (1.4) (CATALYST 4.11 users' manual, 2005).

$$\text{Fit} = \Sigma \text{ mapped hypothesis features} \times W\,[1 - \Sigma\,(\text{disp/tol})^{2}] \ldots\ldots\ldots\; (2.4)$$

where, Σmapped hypothesis features represents the number of pharmacophore features that successfully superimpose (i.e., map or overlap with) corresponding chemical moieties within the fitted compound, *W* is the weight of the corresponding hypothesis feature spheres. This value is fixed to 1.0 in CATALYST-generated models. disp is the distance between the center of a particular pharmacophoric sphere

(feature centroid) and the center of the corresponding superimposed chemical moiety of the fitted compound; tol is the radius of the pharmacophoric feature sphere (known as Tolerance, equals to 1.6 Å by default). $\Sigma(disp/tol)^2$ is the summation of $(disp/tol)^2$ values for all pharmacophoric features that successfully superimpose corresponding chemical functionalities in the fitted compound. (CATALYST 4.11 users' manual, 2005).

The third term, i.e., the configuration cost, penalizes the complexity of the hypothesis. This is a fixed cost, which is equal to the entropy of the hypothesis space. The more the numbers of features (a maximum of five) in a generated hypothesis, the higher is the entropy with subsequent increase in this cost. The overall cost (total cost) of a hypothesis is calculated by summing over the three cost factors. However, error cost is the main contributor to total cost.

CATALYST also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are normally distributed about their mean. Accordingly, the greater the difference from the null hypothesis cost, the more likely that the hypothesis does not reflect a chance correlation. In a successful automatic modeling run, CATALYST ranks the generated models according to their total costs (CATALYST 4.11 users' manual, 2005).

An additional approach to assess the quality of CATALYST-HYPOGEN pharmacophores is to cross-validate them using the Cat-Scramble program implemented in CATALYST. This validation procedure is based on Fisher's randomization test (Fisher, 1966). In this validation test, a 95% confidence level was selected, which instruct CATALYST to generate 19 random spreadsheets by the Cat-Scramble command. Subsequently, CATALYST-HYPOGEN is challenged to use these random spreadsheets to generate hypotheses using exactly the same features and

parameters used in generating the initial unscrambled hypotheses (Krovat *et al*. 2003). Success in generating pharmacophores of comparable cost criteria to those produced by the original unscrambled data reduces the confidence in the training compounds and the unscrambled original pharmacophore models.

2.3.1.6 Clustering of the Generated Pharmacophore Hypotheses

The successful models (336) were clustered into 69 groups utilizing the hierarchical average linkage method available in CATALYST. Subsequently, the highest-ranking representatives, as judged based on their significance r^2-values, were selected to represent their corresponding clusters in subsequent QSAR modeling. (see **Results and discussion,** table 2.4) which shows the statistical criteria of representative pharmacophores including their pharmacophoric features, success criteria and differences from corresponding null hypotheses. The table also shows the corresponding Cat. Scramble confidence levels for each representative pharmacophore

2.3.1.7 QSAR Modeling

A subset of 138 compounds from the total list of inhibitors (1–138, table 2.1) was utilized as a training set for QSAR modeling. However, since it is essential to access the predictive power of the resulting QSAR models on an external set of inhibitors, the remaining 27 molecules (ca. 20% of the dataset) were employed as an external test subset for validating the QSAR models. The test molecules were selected as follows: the collected inhibitors (1–138, table 2.1) were ranked according to their IC_{50} values, and then every fifth compound was selected for the test set starting from the high-

potency end. This selection considers the fact that the test molecules must represent a range of biological activities similar to that of the training set.

The chemical structures of the inhibitors were imported into CERIUS2 as standard 3D single conformer representations in SD format. Subsequently, different descriptor groups were calculated for each compound employing the C2.DESCRIPTOR module of CERIUS2. The calculated descriptors included various simple and valence connectivity indices, electro-topological state indices and other molecular descriptors (e.g., logarithm of partition coefficient, polarizability, dipole moment, molecular volume, molecular weight, molecular surface area, energies of the lowest and highest occupied molecular orbitals, etc.) (CERIUS2, QSAR Users' Manual, 2005). Furthermore, the training compounds were fitted (using the Best-fit option in CATALYST) against the representative pharmacophores (69 models, Table 2.4 in Results and Discussion), and their fit values were added as additional descriptors. The fit value for any compound is obtained automatically via equation 2.4 (CATALYST 4.11 Users' Manual., 2005).

Genetic function approximation (GFA) was employed to search for the best possible QSAR regression equation capable of correlating the variations in biological activities of the training compounds with variations in the generated descriptors, i.e., multiple linear regression modeling (MLR). The fitness function employed herein is based on Friedman's 'lack-of- fit' (LOF) (CERIUS2, QSAR Users' Manual, 2005).

GFA-MLR with large number of poor descriptors; we removed 10% of those showing lowest variance prior to QSAR analysis.

We were obliged to normalize the potencies of the training compounds via division by) (Andrew *et* $\frac{Log(1/IC_{50})}{Mwt}$ their corresponding molecular weights, i.e., ligand efficiency (*al*, 2004), to achieve reasonable self-consistent QSAR models.

Our preliminary diagnostic trials suggested the following optimal GFA parameters: explore linear, quadratic and spline equations at mating and mutation probabilities of 50%; population size = 500; number of genetic iterations = 30,000 and lack-of-fit (LOF) smoothness parameter = 1.0. However, to determine the optimal number of explanatory terms (QSAR descriptors), it was decided to scan and evaluate all possible QSAR models resulting from 5 to 25 explanatory terms.

All QSAR models were validated employing leave one-out cross-validation (r^2_{LOO}), bootstrapping ($r2_{BS}$) and predictive r^2 (r^2_{PRESS}) calculated from the test subsets. The predictive r^2_{PRESS} is defined as:

$$r^2_{PRESS} = SD\text{-}PRESS/SD \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.5)$$

Where SD is the sum of the squared deviations between the biological activities of the test set and the mean activity of the training set molecules, PRESS is the squared deviations between predicted and actual activity values for every molecule in the test set.

2.3.1.8. Receiver Operating Characteristic (ROC) Curve Analysis

QSAR-selected pharmacophore models (i.e., Hypo4/15 and Hypo 6/35) were validated by assessing their abilities to selectively capture diverse ROCK II active compounds from a large testing list of actives and decoys. The testing list was prepared as described by Verdonk and co-workers coworkers (Verdonk *et al.*, 2004; Kirchmair *et al.*, 2008).

Briefly, decoy compounds were selected based on three basic one-dimensional (1D) properties that allow the assessment of distance (D) between two molecules (e.g., i and j): (1) the number of hydrogen-bond donors (NumHBD); (2) number of hydrogen-bond acceptors (NumHBA) and (3) count of nonpolar atoms (NP, defined

as the summation of Cl, F, Br, I, S and C atoms in a particular molecule). For each active compound in the test set, the distance to the nearest other active compound is assessed by their Euclidean Distance (equation. (2.6)):

$$\ldots\ldots(2.6)\ D(i,j)=\sqrt{(NumHBD_i - NumHBD_j)^2+(NumHBA_i - NumHBA_j)^2+(NP_i - NP_j)^2}$$

The minimum distances are then averaged over all active compounds (Dmin). Subsequently, for each active compound in the test set, around 30 decoys were randomly chosen from the ZINC database (Irwin and Shoichet, 2005). The decoys were selected in such a way that they did not exceed Dmin distance from their corresponding active compound.

To diversify active members in the list, we excluded any active compound having zero distance ($D(i,j)$) from other active compound(s) in the test set. Active testing compounds were defined as those possessing ROCK II affinities ranging from 0.01 to 2.2 μM. The test set included 18 active compounds and 642 ZINC decoys.

The test set (660 compounds) was screened by each particular pharmacophore employing the "Best flexible search" option implemented in CATALYST, while the conformational spaces of the compounds were generated employing the "Fast conformation generation option" implemented in CATALYST. Compounds missing one or more features were discarded from the hit list. *In-silico* hits were scored employing their fit values as calculated by Eq. (2.4).

The ROC curve analysis describes the sensitivity (Se or true positive rate, equation 2.7)) for any possible change in the number of selected compounds (n) as a function of (1-Sp). Sp is defined as specificity or true negative rate (equation 2.8) (Kirchmair *et al*., 2008; Triballeau *et al*., 2005).

$$Se = \frac{Number\ of\ Selected\ Actives}{Total\ Number\ of\ Actives} = \frac{TP}{TP + FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.7)$$

$$Sp = \frac{Number\ of\ Discarded\ Inactives}{Total\ Number\ of\ Inactives} = \frac{TN}{TN + FP} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.8)$$

where, TP is the number of active compounds captured by the virtual screening method (true positives), FN is the number of active compounds discarded by the virtual screening method, TN is the number of discarded decoys (presumably inactives), while FP is the number of captured decoys (presumably inactive) (Kirchmair *et al*., 2008; Triballeau *et al*., 2005).

If all molecules scored by a virtual screening (VS) protocol with sufficient discriminatory power are ranked according to their score (i.e., fit values), starting with the best-scored molecule and ending with the molecule that got the lowest score, most of the actives will have a higher score than the decoys. Since some of the actives will be scored lower than decoys, an overlap between the distribution of active molecules and decoys will occur, which will lead to the prediction of false positives and false negatives (Kirchmair *et al*., 2008; Triballeau *et al*., 2005). The selection of one score

value as a threshold strongly influences the ratio of actives to decoys and therefore the validation of a VS method. The ROC curve method avoids the selection of a threshold by considering all Se and Sp pairs for each score threshold (Triballeau *et al*., 2005). A ROC curve is plotted by setting the score of the active molecule as the first threshold. Afterwards, the number of decoys within this cutoff is counted and the corresponding Se and Sp pair is calculated. This calculation is repeated for the active molecule with the second highest score and so forth, until the scores of all actives are considered as selection thresholds.

The ROC curve representing ideal distributions, where no overlap between the scores of active molecules and decoys exists, proceeds from the origin to the upper-left corner until all the actives are retrieved and Se reaches the value of 1. In contrast to that, the ROC curve for a set of actives and decoys with randomly distributed scores tends towards the Se = 1-Sp line asymptotically with increasing number of actives and decoys(Kirchmair *et al*., 2008; Triballeau *et al*., 2005). The success of a particular virtual screening workflow can be judged from the following criteria:

1) Area under the ROC curve (AUC) (Triballeau *et al*., 2005). In an optimal ROC curve an AUC value of 1 is obtained; however, random distributions cause an AUC value of 0.5. Virtual screening that performs better than a random discrimination of actives and decoys retrieve an AUC value between 0.5 and 1, whereas an AUC value lower than 0.5 represents the unfavorable case of a virtual screening method that has a higher probability to assign the best scores to decoys than to actives (Kirchmair *et al*., 2008; Triballeau *et al*., 2005).

2) Overall accuracy (ACC): describes the percentage of correctly classified molecules by the screening protocol (equation 2.9). Testing compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) (Gao *et al*., 1999; Jacobsson *et al*., 2003; Kirchmair *et al*., 2008).

$$ACC = \frac{TP + TN}{N} = \frac{A}{N}.Se + \left(1 - \frac{A}{N}\right).Sp \text{ ....(2.9)}$$

where, N is the total number of compounds in the testing database, A is the number of true actives in the testing database.

3) Overall specificity (SPC): describes the percentage of discarded inactives by the particular virtual screening workflow. Inactive test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values (Gao *et al*., 1999; Jacobsson *et al*., 2003; Kirchmair *et al*., 2008).

4) Overall true positive rate (TPR or overall sensitivity): describes the fraction percentage of captured actives from the total number of actives. Active test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values.

5) Overall false negative rate (FNR or overall percentage of discarded actives): describes the fraction percentage of active compounds discarded by the virtual screening method. Discarded active test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values (Abu Khalaf *et al*., 2010).

2.3.1.9 *In Silico* Screening for New ROCK II inhibitors:

Hypo4/15 and Hypo6/35 were employed as 3D search queries to screen the National Cancer Institute (NCI) 3D flexible structural database. The screening was done employing the "Best Flexible Database Search" option implemented within CATALYST. Captured hits were filtered according to Lipinski's and Veber's rules (Lipinski *et al.,* 2001; Veber *et al*., 2002). Remaining hits were fitted against the two pharmacophores using the "best fit" option within CATALYST. The fit values together with the relevant molecular descriptors of each hit were substituted in the optimal QSAR equation (2.10). The highest ranking molecules based on QSAR predictions were acquired and tested *in vitro*. Table 2.8 shows active hits and their QSAR-predictions and experimental bioactivities.

2.3.1.10. *In vitro* Experimental studies

Materials

All of the chemicals used in these experiments were of reagent grade and obtained from commercial suppliers. NCI samples were kindly provided by the national cancer institute. Rho-kinase drug discovery kit was purchased from Cyclex (Japan), the standard inhibitor Y-27632, water and DMSO for bioanalysis were all purchased form Sigma-Aldrich (USA).

Preparation of hit compounds for *In vitro* assay

The tested compounds were provided as dry powders in variable quantities (5 mg to 10 mg). They were initially dissolved in DMSO to give stock solutions of fixed concentrations. Subsequently, they were diluted to the required concentrations with deionized water for bioassay.

Quantification of Rho-kinase activity in a spectrophotometric assay

The activity of the in silico hits were quantified by rho-kinase drug discovery kit (Cyclex, Japan). The 96-well plate of the kit is pre-coated with recombinant C terminus of the myosin-binding subunit of myosin phosphatase (MBS) phosphorylated by rho-kinase (at threonine-696). The detector antibody is conjugated to horseradish peroxidase and specifically detects the phosphorylated form of MBS.

To perform the assay, the tested hits were first diluted with the kinase buffer provided with the kit. The hits' solutions were then pipetted into assay wells to yield final concentrations of 0.1 to 10 μM. Subsequently, rho-kinase was added to each well as aqueous solution (0.01 units in 10 μL) and allowed to phosphorylate the bound substrate in the presence of Mg^{2+} and ATP. The amount of phosphorylated substrate was measured by binding it with the detector antibody conjugate to horseradish peroxidase, which then catalyzes the conversion of the chromogenic substrate tetra-methylbenzidine (TMB) from a colorless solution to a blue solution. The color is quantified by spectrophotometry and reflects the relative activity of rho-kinase. Samples' absorbances were determined at a wavelength of 650 nm using a plate reader (Bio-Tek instruments ELx 800, USA). Inhibition of Rho-kinase was calculated as

percent activity of the uninhibited Rho-kinase enzyme control. Y-27632 was tested as positive control. Negative controls were prepared by adding the enzyme to the reaction with the kinase reaction buffer only (Cyclex Rho-Kinase Assay Kit (Cat# CY-1160) Users' Manual, 2009).

2.4. Results and Discussion

Different binding hypotheses were generated for a series of 138 ROCK II inhibitors (Table 2.1). Six training subsets were selected from the collection (Table 2.2). Each subset consisted of inhibitors of wide structural diversity. The biological activity in the training subsets spanned around 3.5 order of magnitude. Genetic algorithm and multiple linear regression statistical analysis were subsequently employed to select an optimal combination of complementary pharmacophores capable of explaining bioactivity variations among all inhibitors.

2.4.1 Data Mining and Conformational Coverage

The literature was surveyed to collect many reported structurally diverse ROCK II inhibitors (**1-138**, table 2.1). The 2D structures of the inhibitors were imported into CATALYST and converted automatically into 3D single conformer representations. The structures were used as starting points for conformational analyses and in the determination of various molecular descriptors for QSAR modeling.

The conformational space of each inhibitor was extensively sampled utilizing the poling algorithm employed within the CONFIRM module of CATALYST CATALYST 4.11 users' manual, 2005. Conformational coverage was performed employing the "Best" module to ensure extensive sampling of conformational space.

Pharmacophore generation and pharmacophore-based search procedures are known for their sensitivity to inadequate conformational sampling within the training compounds prompting our extensive conformational sampling (Abu-Hammad *et al*, 2008).

2.4.2 Exploration of ROCK II Pharmacophoric Space

The fact that we have an informative list of 138 ROCKII inhibitors of evenly spread bioactivities over more than 3.5 orders of magnitude, prompted us to employ CATALYST-HYPOGEN to identify possible pharmacophoric binding modes assumed by ROCKII inhibitors (Takami *et al*, 2004, Iwakubo *et al*, 2007, a&b).

HYPOGEN implements an optimization algorithm that evaluates large number of potential binding models for a particular target through fine perturbations to hypotheses that survived the constructive and subtractive phases of the modeling algorithm (see section **2.3.1.4 Pharmacophoric Hypotheses Generation** in **Experimental**) (Li *et al*., 2000) The extent of the evaluated pharmacophoric space is reflected by the configuration (Config.) cost calculated for each modeling run. It is generally recommended that the Config. cost of any HYPOGEN run not to exceed 17 (corresponding to 2^{17} hypotheses to be assessed by CATALYST) to guarantee thorough analysis of all models (Sutter *et al.,* 2000). The size of the investigated pharmacophoric space is a function of training compounds, selected input chemical features and other CATALYST control parameters (Sutter *et al.,* 2000).

Restricting the extent of explored pharmacophoric space should improve the efficiency of optimization via allowing effective evaluation of limited number of pharmacophoric models. On the other hand, rigorous restrictions imposed on the

pharmacophoric space might reduce the possibility of discovering optimal pharmacophoric hypotheses, as they might occur outside the “boundaries” of the pharmacophoric space.

Therefore, we decided to explore the pharmacophoric space of ROCKII inhibitors under reasonably imposed "boundaries" through 36 HYPOGEN automatic runs and employing six carefully selected training subsets: subsets **I-VI** (Table 2.2). The training compounds in these subsets were selected in such away to guarantee maximal 3D diversity and continuous bioactivity spread over more than 3.5 logarithmic cycles. We gave special emphasis to the 3D diversity of the most active compounds in each training subset (Table 2.2) because of their significant influence on the extent of the evaluated pharmacophoric space during the constructive phase of HYPOGEN algorithm (see section **2.3.1.4 Pharmacophoric Hypotheses Generation in Experimental**) (Li *et al*, 2000).

Guided by our rationally restricted pharmacophoric exploration concept, we restricted the software to explore pharmacophoric models incorporating from zero to one PosIon feature, from zero to three HBA, HBD, Hbic, and RingArom features instead of the default range of zero to five, (Table 2.3). Furthermore, we instructed HYPOGEN to explore only 4- and 5-featured pharmacophores, i.e., ignore models of lesser number of features in order to further narrow the investigated pharmacophoric space and to represent the feature-rich nature of known ROCK II ligands (Table 2.3).

In each run, the resulting binding hypotheses were automatically ranked according to their corresponding "total cost" value, which is defined as the sum of error cost, weight cost and configuration cost (see section **2.3.1.5 Assessment of the Generated Hypotheses in Experimental**) Error cost provides the highest contribution to total

cost and it is directly related to the capacity of the particular pharmacophore as 3D-QSAR model, i.e., in correlating the molecular structures to the corresponding biological responses. HYPOGEN also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are normally distributed about their mean. Accordingly, the greater the difference from the null hypothesis cost (residual cost, table 2.4) the more likely that the hypothesis does not reflect a chance correlation (Bersuker *et al*. 2000; CATALYST 4.11 users' manual, 2005; Kurogi and Güner, 2001; Li *et al*., 2000; Poptodorov *et al*. 2006; Sutter *et al*. 2000).

An additional validation technique based on Fischer's randomization test (Fischer *et al*, 1966) was recently introduced into CATALYST: Cat. Scramble (CATALYST 4.11 users' manual, 2005). In this test the biological data and the corresponding structures are scrambled several times and the software is challenged to generate pharmacophoric models from the randomized data. The confidence in the parent hypotheses (i.e., generated from unscrambled data) is lowered proportional to the number of times the software succeeds in generating binding hypotheses from scrambled data of apparently better cost criteria than the parent hypotheses (see section **2.3.1.5 Assessment of the Generated Hypotheses in Experimental**). Fortunately, all generated pharmacophores illustrated ≥ 90% Cat.Scramble significance. Eventually, 336 pharmacophore models emerged from 36 automatic HYPOGEN runs. These were subsequently clustered and their best representatives (69 models, Table 2.4, see section **2.3.1.6 Clustering of the Generated Pharmacophore Hypotheses**) were used in subsequent QSAR modeling. Clearly from table 2.4, the representative models shared comparable features and excellent statistical success criteria. Emergence of several statistically comparable

pharmacophore models suggests the ability of ROCKII ligands to assume multiple pharmacophoric binding modes within the binding pocket. Therefore, it is quite challenging to select any particular pharmacophore hypothesis as a sole representative of the binding process.

Table 2.4: Success criteria of representative pharmacophoric hypotheses (cluster centers).

RU N[a]	Hypotheses[b]	Features	Cost				R[d]	Global R[e]	F-statisti[f]	Cat. Scrambl
			Config.	Tot	Null	Residual				
1	3	HBDN, Hbic, 2xRingArom	14.1	108.	125.	16.9	0.85	0.0	14.2	95%
	10	HBA, HBDN, Hbic,	14.1	110.	125.	14.8	0.82	0.0	4.6	95%
2	4	HBA, HBDN, Hbic,	10.9	102.	125.	23	0.9	0.1	30.3	95%
	8	HBA, HBDN, Hbic,	10.9	106.	125.	18.8	0.83	0.2	38.6	95%
	10	3xHBDN, Hbic,	10.9	106.	125.	18.8	0.83	0.1	30.6	95%
3	3	HBDN, 3xHbic, RingArom	9.2	110.	125.	14.6	0.78	0.0	9	95%
	4	2xHBDN, 3xHbic	9.2	112.	125.	12.6	0.77	0.1	20.7	90%
	6	HBA, HBDN,3xHbic	9.2	113.	125.	12.2	0.76	0.1	21.4	90%
5	4	2xHBDN, Hbic, RingArom	12.5	108.	125.	16.9	0.83	0.2	49.6	90%
	8	2xHBDN, 2x Hbic	12.5	109.	125.	15.6	0.83	0.2	37.8	90%
6	9	HBA, 2xHBDN, Hbic	9.6	106.	125.	18.6	0.82	0.2	53.1	95%
7	2	2xHBDN, 3xHbic	7.4	104.	125.	20.6	0.83	0.1	17.3	90%
9	7	2xHbic, RingArom, PosIon	17	109.	186.	77.4	0.97	0.2	41.5	95%
	8	Hbic, 2xRingArom, PosIon	17	109.	186.	77.5	0.97	0.2	34.7	95%
10	2	HBA, Hbic, RingArom,	15.1	104.	186.	81.8	0.98	0.2	40	95%
	3	HBA, Hbic, RingArom,	15.1	104.	186.	81.7	0.98	0.2	47.9	95%
	7	HBA, 2xHbic, PosIon	15.1	106.	186.	80.3	0.97	0.2	44.6	95%
	9	HBD, 2xRingArom, PosIon	15.1	106.	186.	80	0.97	0.2	39.4	95%
	10	HBA, 2xRingArom, PosIon	15.1	106.	186.	79.8	0.97	0.2	40.2	95%
12	2	2xHBD, 2xHbic, PosIon	9.7	107.	186.	78.9	0.94	0.2	48	95%
14	2	2xHBD, Hbic, PosIon	12.2	95.7	149.	53.7	0.95	0.2	39.7	95%
	3	HBA, 2xHbic, PosIon	12.2	98.7	149.	50.7	0.93	0.2	46.6	95%
	7	HBA, 2xRingArom, PosIon	12.2	100.	149.	48.9	0.91	0.2	41.8	95%
15	**4**	**HBD, 2xHbic, RingArom,**	**12.7**	**93.5**	**149.**	**55.9**	**0.96**	**0.2**	**53.3**	**95%**
	7	HBD, 3xHbic, PosIon	12.7	95.8	149.	53.6	0.96	0.2	38.8	95%
16	5	HBA, HBD, 3xHbic	3.7	102	149.	47.4	0.83	0.1	25.6	95%
	6	HBA, HBD, 2xHbic, PosIor	3.7	102.	149.	46.8	0.83	0.3	61.8	95%
17	3	2xHBD, Hbic, , PosIon	16.1	94.8	113.	19.1	0.95	0.1	16.8	95%
	7	HBA, HBD, Hbic, PosIon	16.1	95.1	113.	18.9	0.95	0.2	52.9	95%
	8	HBA,2xHBD, PosIon	16.1	95.2	113.	18.7	0.95	0.0	12.6	95%
18	1	HBA, HBD, Hbic, PosIon	13.8	92.1	113.	21.8	0.96	0.0	14.1	95%
	4	HBA, HBD, Hbic, PosIon	13.8	92.1	113.	21.8	0.96	0.0	10.2	95%
	7	HBA, HBD, RingArom,	13.8	92.4	113.	21.5	0.96	0.1	30.9	95%
	8	HBA, HBD, RingArom,	13.8	92.6	113.	21.3	0.96	0.3	65.8	95%
19	1	HBD, 2xHbic, RingArom,	13.7	94	113.	0	0.94	0.2	54.9	95%
	10	HBD, 2xHbic, RingArom,	13.7	89.6	113.	19.9	0.92	0.2	38.5	95%
20	4	HBA, HBD, 2xHbic, PosIor	8.6	89.7	113.	24.3	0.93	0.2	36.4	95%
	6	HBA, HBD, 2xHbic, PosIor	8.6	94.2	113.	0	0.87	0.1	21.3	95%
	9	HBA, HBD, 2xHbic, PosIor	8.6	96.4	113.	24.2	0.85	0.2	51.5	95%
22	3	2xHBD, Hbic, PosIon	11.9	91.5	113.	19.8	0.94	0.0	6.2	95%
23	8	2xHBD, 2xHbic, PosIon	11.4	97	113.	17.5	0.89	0.2	38.1	95%
24	3	2xHBD, 2xHbic, PosIon	2.6	98.5	113.	22.4	0.77	0.0	6.6	95%
	6	2xHBD, 3xHbic	2.6	109. 6	113. 9	16.9	0.59	0.0 1	1	95%

Table 2.4: Success criteria of representative pharmacophoric hypotheses (cluster centers).

RUN [a]	Hypotheses[b]	Features	Cost				R^d	Global R^e	F-statistic[f]	Cat. Scrambl
			Config.	**Tota**	**Null**	**Residual[c]**				
25	1	HBD, 2xHbic, PosIon	15.8	93.7	161.	67.8	0.99	0.2	34.6	95%
	5	HBA, HBD, RingArom, PosIon	15.8	96.2	161.	65.4	0.97	0.29	54.1	95%
	8	HBD, Hbic, RingArom, PosIon	15.8	96.8	161.	64.8	0.97	0.19	31.7	95%
26	2	HBA, Hbic, RingArom, PosIon	12.9	92.4	161.	69.2	0.98	0.28	51.7	95%
	9	HBA, HBD, RingArom, PosIon	12.9	94.6	161.	67	0.96	0.25	45.5	95%
	10	2xHBD, Hbic, PosIon	12.9	94.8	161.	66.8	0.96	0.21	36.3	95%
27	9	HBD, 2xHbic, RingArom,	14.1	96.3	161.	65.2	0.97	0.26	48.7	95%
28	8	HBA, HBD, 2xHbic, PosIon	6	98.3	161.	63.3	0.91	0.31	62.1	95%
29	3	HBAF, 2xRingArom, PosIon	17.8	91	117.	26.8	0.94	0.28	52	95%
	5	2xHBAF, RingArom, PosIon	17.8	91.1	117.	26.7	0.92	0.11	16.9	95%
	8	2xHBAF, RingArom, PosIon	17.8	91.9	117.	25.9	0.91	0.09	13.9	95%
	9	HBDN, 2xRingArom, PosIon	17.8	92	117.	25.8	0.91	0.23	39.6	95%
30	2	3xHBAF, Hbic	15.8	86.8	117.	31	0.95	0.06	8.5	95%
	4	HBAF, HBDN, Hbic ,	15.8	86.8	117.	30.9	0.95	0.21	35.4	95%
	7	2xHBAF, Hbic, PosIon	15.8	87.1	117.	30.6	0.94	0.11	16.9	95%
	10	HBAF, HBDN, RingArom,	15.8	87.5	117.	30.3	0.94	0.26	47.8	95%
31	3	HBAF, 2xHbic, RingArom,	15.5	89	117.	28.8	0.92	0.31	60.4	95%
	8	HBDN, 2xHbic, RingArom,	15.5	91	117.	26.7	0.9	0.22	39.2	95%
	10	2xHBDN, 2xHbic, PosIon	15.5	92	117.	25.8	0.9	0.23	40.7	95%
32	1	HBAF, HBDN, 2xHbic , PosIon	11.4	88.1	117.	29.7	0.89	0.28	51.9	95%
	8	2xHBAF, 2xHbic, PosIon	11.4	91.6	117.	26.1	0.86	0.07	10.3	95%
33	1	HBAF, Hbic, RingArom, PosIon	13.3	80.4	126.	45.8	0.98	0.31	60.6	95%
34	5	HBAF, Hbic, RingArom, PosIon	24.9	92.2	126.	33.9	0.98	0.3	59	95%
35	**6**	**HBAF, Hbic, HbicArom,**	**11.4**	**80.5**	**126.**	**45.6**	**0.97**	**0.23**	**40.6**	**95%**
	10	HBAF, Hbic, HbicArom, PosIon	11.4	81.3	126. 2	44.8	0.96	0.23	41.1	95%

[a]Correspond to runs in Table 2.3, [b]High ranking representative hypotheses (in their corresponding clusters, see section **2.3.1.6 Clustering of the Generated Pharmacophore Hypotheses**). [c]Difference between total cost and the cost of the corresponding null hypotheses. [d]Correlation coefficients between pharmacophore-based bioactivity estimates (calculated from equation (2.3) and bioactivities of corresponding training compound (subsets in table 2.2). [e]Correlation coefficients between pharmacophore-based bioactivity estimates and bioactivities of all collected compounds. [f]Fisher statistic calculated based on the linear regression between the fit values of all collected inhibitors (**1-138**, table 1.1) against pharmacophore hypothesis (employing the "best fit" option and equation (2.4) and their respective anti-ROCK II bioactivities (log ($1/IC_{50}$) values). [g]ranking of hypotheses is as generated by CATALYST in each automatic run. [h]Bolded pharmacophores appeared in the best QSAR equations.

2.4.3 QSAR Modeling

Despite that pharmacophoric hypotheses provide excellent insights into ligand-macromolecule recognition and can be used to mine for new biologically interesting scaffolds, their predictive value as 3D-QSAR models is limited by steric shielding and bioactivity-enhancing or reducing auxiliary groups (Taha *et al.*, 2008, a,b&c; Taha *et al.*, 2007; Taha *et al.*, 2010; Al-masri *et al*; 2008; Al-Nadaf *et al*; 2010; Abu-Hammad *et al*, 2009; Abu Khalaf *et al*, 2010; Al-Sha'er and Taha, 2010, a&b; Abdula *et al*, 2011; Abu Khalaf *et al*, 2011). This point combined with the fact that pharmacophore modeling of ROCKII inhibitors furnished several binding hypotheses of comparable success criteria prompted us to employ classical QSAR analysis to search for the best combination of pharmacophore(s) and other 2D descriptors capable of explaining bioactivity variation across the whole list of collected inhibitors (**1-138**, table 2.1). We employed genetic function approximation and multiple linear regression QSAR (GFA-MLR-QSAR) analysis to search for an optimal QSAR equation(s).

The fit values obtained by mapping representative hypotheses (69 models) against collected ROCK II inhibitors (**1-138**, table 2.1) were enrolled, together with around 100 other physicochemical descriptors, as independent variables (genes) in GFA-MLR-QSAR analysis (see section **2.3.1.7 QSAR Modeling in Experimental**) (Taha *et al.*, 2008, a,b&c; Taha *et al.*, 2007; Taha *et al.*, 2010; Al-masri *et al*; 2008; Al-Nadaf *et al*; 2010; Abu-Hammad *et al*, 2009; Abu Khalaf *et al*, 2010; Al-Sha'er *et al*, 2010, a&b; Abdula *et al*, 2011; Abu Khalaf *et al*, 2011, CERIUS2 QSAR Users' Manual, 2005; Ramsey and Schafer, 1997) However, since it is essential to access the predictive power of the resulting QSAR models on an external set of inhibitors, we randomly selected 27 molecules (marked with double asterisks in table 2.1) and

employed them as external test molecules for validating the QSAR models (r^2_{PRESS}). Moreover, all QSAR models were cross-validated automatically using the leave-one-out cross-validation in CERIUS2 (CERIUS2 QSAR Users' Manual, 2005; Ramsey and Schafer, 1997).

Unfortunately, all our attempts to achieve self-consistent and predictive QSAR models were futile prompting us to evaluate an alternative modeling strategy, namely, to employ ligand efficiency [log(IC_{50})/MW] as the response variable instead of activity [log(IC_{50})]. The fact that many training compounds exhibited similar potencies despite their wide structural variations (e.g., 6 training compounds illustrated IC_{50} values of 0.1 μM and 39 training compounds illustrated IC_{50} values of 35 μM) prompted us to maximize bioactivity variance through division (normalization) over molecular weights. This novel strategy proved successful in achieving self-consistent QSAR models (Al-Nadaf *et al*, 2010).

Equation (2.10) shows the details of the optimal QSAR model. Figure 2.2 shows the corresponding scatter plots of experimental versus estimated bioactivities for the training and testing inhibitors.

$$\frac{Log(1/IC_{50})}{MW} = 0.24 + 1.07\text{x}10^{-2}(\text{Hypo6/35}) + 5.2\text{x}10^{-3}(\text{Hypo4/15}) + 2.04\text{x}10^{-2}(\text{AtypeN72}) - 5.1\text{x}10^{-3}(\text{AtypeH53}) - 7.4\text{x}10^{-3}(\text{AtypeC24}) - 5.2\text{x}10^{-2}(\text{LUMO}) - 0.27(\text{JX}) - 2.2\ \text{x}10^{-4}\ \text{Wiener} + 0.28\chi^0$$

$r_{111} = 0.823$, F-statistic $= 18.75$, $r^2_{BS} = 0.604$, $r^2_{LOO} = 0.679$, $r^2_{PRESS(27)} = 0.535$………........... (2.10)

where, r_{111} is the correlation coefficient against 111 training compounds, r^2_{LOO} is the leave-one-out correlation coefficient, r^2_{BS} is the bootstrapping regression coefficient and r^2_{PRESS} is the predictive r^2 determined for the 16 test compounds (CERIUS2 QSAR Users' Manual, 2005; Ramsey and Schafer, 1997).

Hypo4/15 and Hypo6/35 represent the fit values of the training compounds against these two pharmacophores (bolded models in table 2.4 and figures 2.3 and 2.4) as calculated from equation (2.4). Atype descriptors are atom-type-based AlogP thermodynamic descriptors: AtypeN72 encodes for the number of amidic NH, AtypeC24 encodes for the number of aromatic CH atoms (excluding symmetrical ones), AtypeH53 encodes for the substitution patterns of alicyclic rings with 1,4-disubstituted cyclohexyls are given the highest values while 1,3-disubstituted analogues are assigned the least values.

LUMO is the energy of the lowest unoccupied molecular orbital calculated by semiempirical quantum mechanical method (MOPAC). JX is a Balaban index related to the shape and atomic electronegativities of a particular compound. Wiener is the sum of the chemical bonds existing between all pairs of heavy atoms in the molecule, and χ^0 is the zero order chi connectivity index. Table 4.1 in Appendix shows the values of the different descriptors in equation (2.10) for modeled compounds (Table 2.1).

Emergence of two orthogonal pharmacophoric models, i.e., Hypo4/15 and Hypo6/35 of cross-correlation $r^2 \leq 0.28$, in equation (2.10) suggests they represent two complementary binding modes accessible to ligands within the binding pocket of

ROCK II, i.e., one of the pharmacophores explains the bioactivities of some training inhibitors while the other explains the remaining inhibitors. Similar conclusions were reached about the binding pockets of other targets based on QSAR analysis (Taha *et al.*, 2008, a,b&c; Taha *et al.*, 2007; Taha *et al.*, 2010; Al-masri *et al*; 2008; Al-Nadaf *et al*; 2010; Abu-Hammad *et al*, 2009; Abu Khalaf *et al*, 2010; Al-Sha'er and Taha, 2010, a&b; Abdula *et al*, 2011; Abu Khalaf *et al*, 2011).Figures 2.6C and 2.7C show Hypo6/35 and Hypo4/15 and how they map **107** (IC_{50} = 0.009 μM) and **98** (IC_{50} = 0.02 μM), respectively. The X, Y, and Z coordinates of the two pharmacophores are given in table 2.5.

Emergence of AtypeC24 in conjunction with negative slope suggests that ROCK II inhibition coincides with increased aromatic substitution (i.e., lesser aromatic CHs), which is probably related to increased ligands' attractive interactions with binding site residues via aromatic substituents. On the other hand, appearance of AtypeN72 combined with positive regression coefficient suggests that anti-ROCKII bioactivity is proportional to the number of amidic NHs in the ligand. This trend is probably because of hydrogen-bonding interactions within the ligand-receptor complex involving ligands' amidic nitrogens.

However, emergence of AtypeH53 in associate with negative slope in QSAR equation (2.10) seems to encode for a complex SAR relationship: ligands possessing 1,3-disubstituted cyclohexyl (or cyclopentyl) cores seem to better fit the binding site compared to other substitution patterns, e.g., 1,4-disubstituted cyclohexyl-based cores.

Furthermore, emergence of LUMO in equation (2.10) combined with negative slope suggests that ligand/ROCK II affinity favors electrophilic ligands probably due to π-

stacking with certain electron-rich aromatic centers in the binding pocket, probably the aromatic rings of Phe103 and Phe384 (Figure 2.4A).

Finally, emergence of Wiener and $\square^{0}$ connectivity indices in equation (2.10) illustrates certain role played by the ligands' topology in the binding process. However, despite their predictive significance, the information content of topological descriptors is quite obscure.

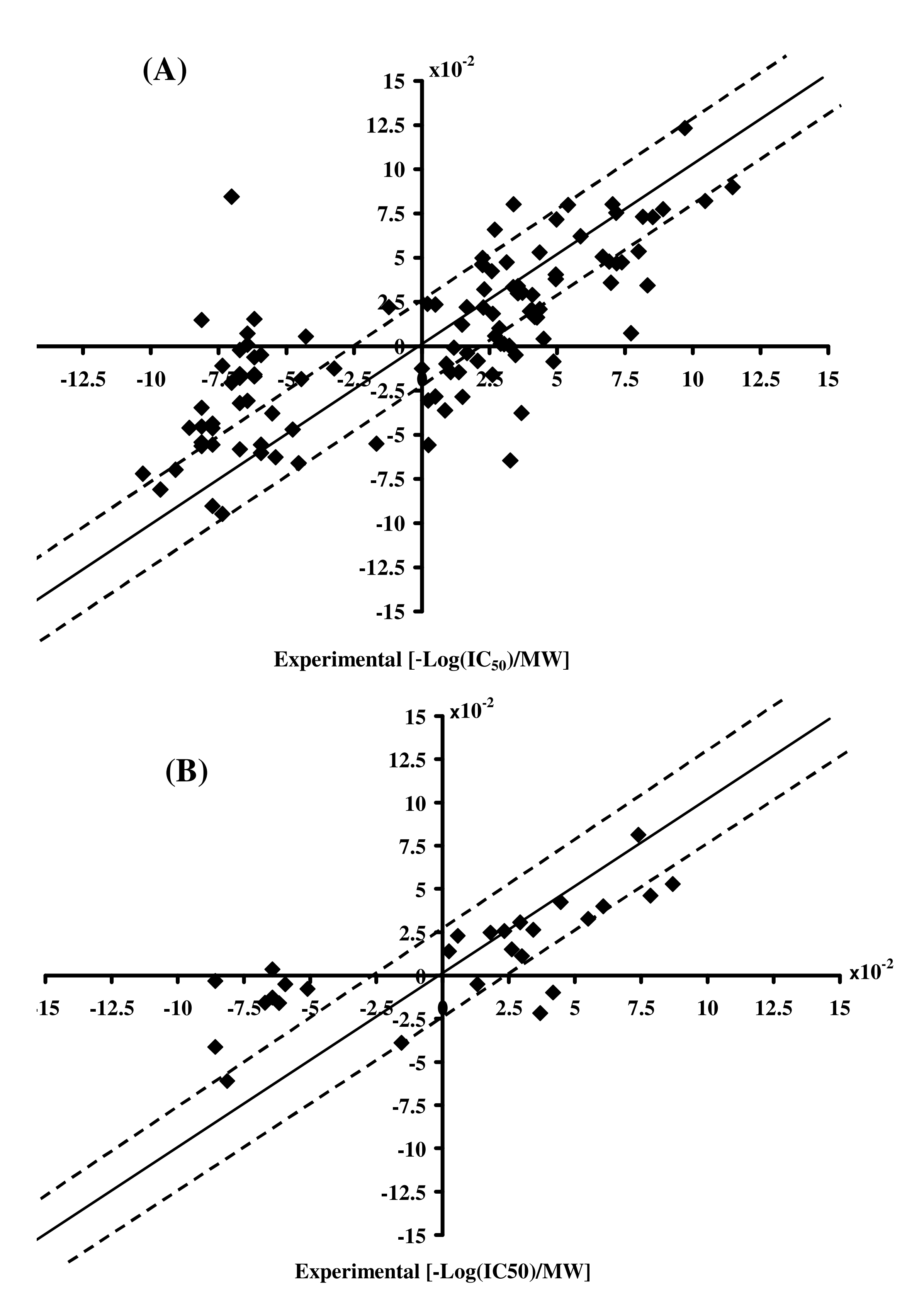


Figure 2.2: Experimental versus **(A)** fitted (111 compounds, $r2_{LOO}$= 0.678), and **(B)** predicted (27 compounds, $r2_{PRESS}$= 0.535) bioactivities calculated from the best QSAR model equation (2.10). The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively, whereas the dotted lines indicate the $2.5x10^{-2}$ (-logIC_{50})/MW error margins.

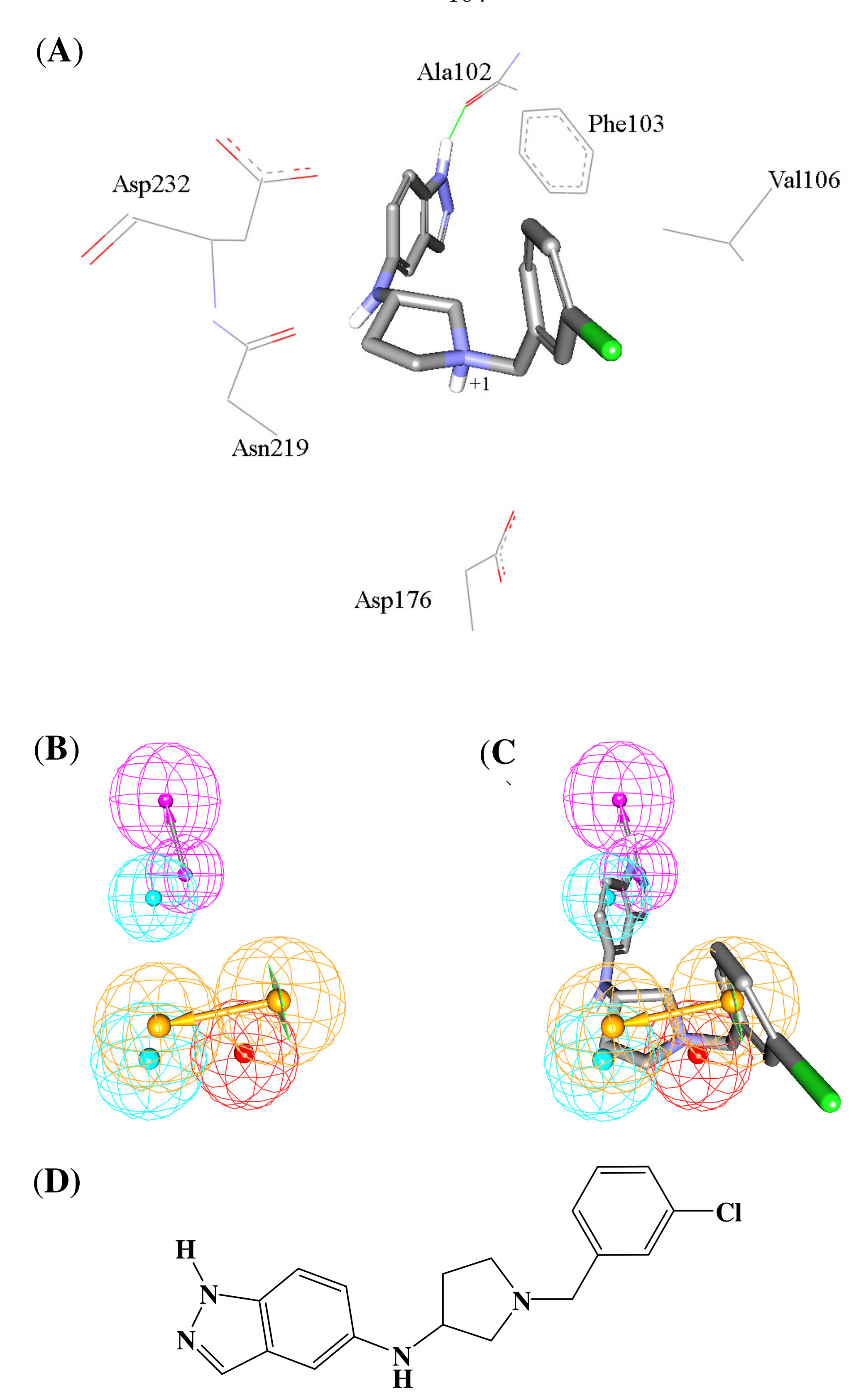


Figure 2.3 (A) Docked structure of training compound **98** (IC_{50} = 0.02 µM, table 1.1) into ROCK II (PDB code: 2H9V, resolution 1.85 Å). (**B**) Pharmacophoric features of Hypo4/15: HBD as pink vectored spheres, Hbic as light blue spheres, RingArom as vectored orange spheres, PosIon as red spheres. **(C)** Hypo4/15 fitted against **98**, **(D)** Chemical structure of **98**.

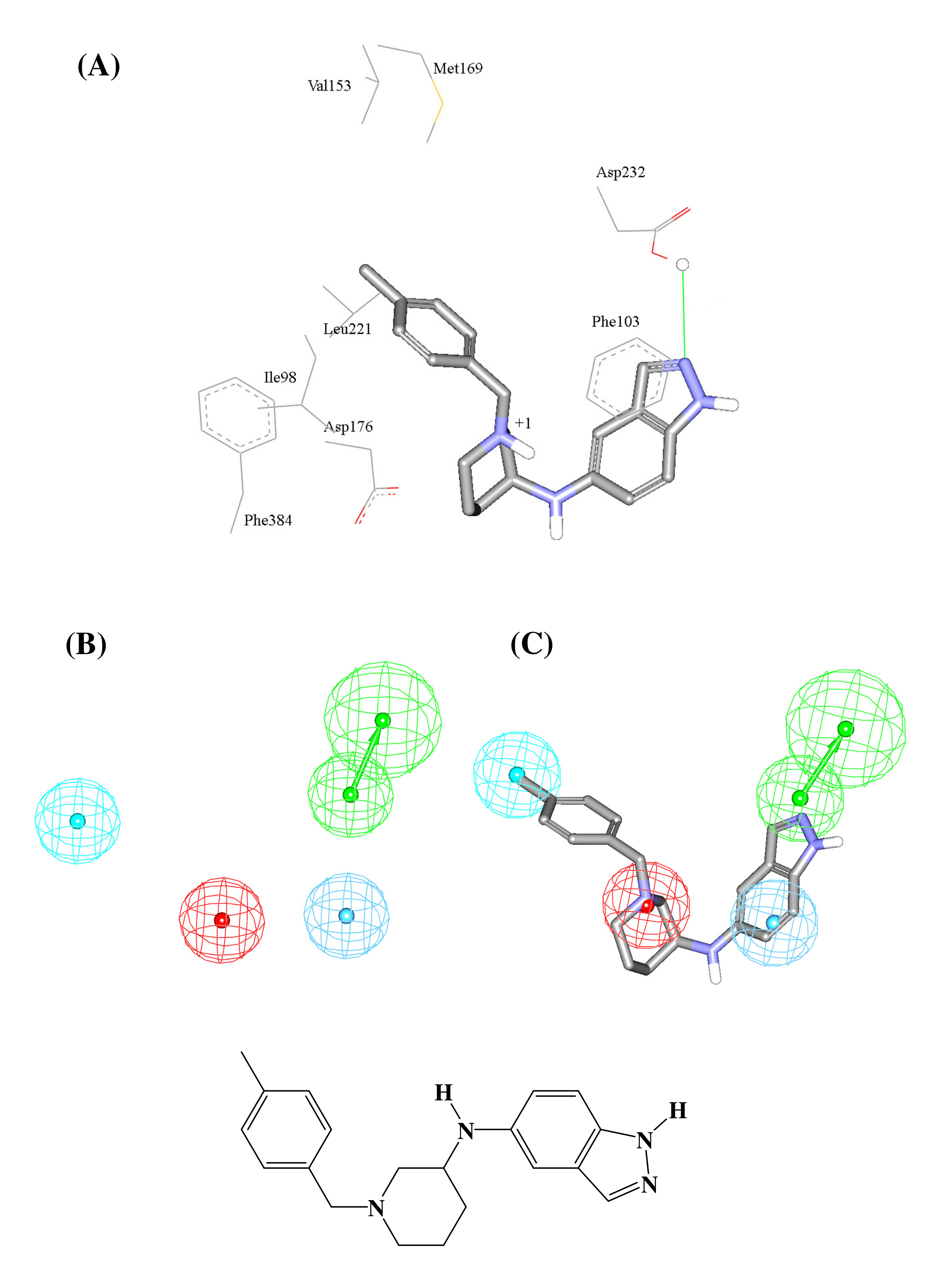


Figure 2.4: **(A)** Docked structure of training compound **107** (IC_{50} = 0.009 µM, table 1.1) into ROCK II (PDB code: 2H9V, resolution 1.85 Å) binding pocket. (**B**) Pharmacophoric features of Hypo6/35: HBA as green vectored spheres, Hbic as light blue spheres, HbicArom as dark blue spheres, PosIon as red spheres **(C)** Hypo6/35 fitted against **107**, **(D)** Chemical structure of **107**.

Table 2.5: Pharmacophoric features and corresponding weights, tolerances and 3D coordinates of Hypo4/15, Hypo6/35.

Model	definition		Chemical Features						
			HBD		**RingArom**		**Hbic**	**Hbic**	**PosIon**
Hypo4/15[a]	Weights		1.93610		1.93610		1.93610	1.93610	1.93610
	Tolerances		1.60	2.20	1.60	2.20	1.60	1.60	1.60
	Coordinate	X	4.21	4.77	-3.72	-1.99	2.94	-0.84	-1.77
		Y	1.57	1.00	-1.13	-3.49	-0.67	-3.10	-0.43
		Z	-3.13	-6.03	-0.36	0.32	-2.35	1.72	1.88
			HBAN		**Hbic**	**HbicArom**	**PosIon**		
Hypo6/35[b]	Weights		2.22164		2.22164	2.22164	2.22164		
	Tolerances		1.60	2.20	1.60	1.60	1.60		
	Coordinate	X	0.78	2.13	0.38	-2.03	-0.57		
		Y	-5.11	-7.59	-0.38	-4.73	-0.37		
		Z	4.49	5.53	4.40	-5.40	-0.15		

[a]Hypo4/15: the 4th pharmacophore hypothesis generated in the 15th HYPOGEN run (table 2.4).
[b]Hypo6/35: the 6th pharmacophore hypothesis generated in the 35th HYPOGEN run (table 2.4).

2.4.4 Receiver Operating Characteristic (ROC) Curve Analysis and Shape Constraints

To further validate the resulting models (both QSAR and pharmacophores), we subjected our QSAR-selected pharmacophores to receiver-operating curve (ROC) analysis. In ROC analysis, the ability of a particular pharmacophore model to correctly classify a list of compounds as actives or inactives is indicated by the area under the curve (AUC) of the corresponding ROC together with other parameters, namely, overall accuracy, overall specificity, overall true positive rate and overall false negative rate (see Section **2.3.1.8** under **Experimental** for more details) (Verdonk *et al*, 2004 , Kirchmair *et al*, 2008 , Irwin *et al*, 2005 , Triballeau *et al*, 2005).

Table 2.6 and figure 2.5 show the ROC results of our QSAR-selected pharmacophores. Hypo4/15 and Hypo6/35 illustrated excellent overall performances with AUC values of 99.5% and 99.4% respectively. This is not unexpected, as the presence of positive ionizable features in both models should significantly enhance their selectivities.

Table 2.6: ROC curve analysis criteria for QSAR-selected pharmacophores and their sterically-refined versions.

Pharmacophore Model	ROC[a]-AUC%[b]	ACC%[c]	SPC%[d]	TPR%[e]	FNR%[f]
Hypo4/15	99.5	97.3	99.7	11.1	0.3
Hypo6/35	99.4	97.3	99.2	27.8	0.8

[a]**ROC**: receiver operating characteristic curve.
[b]**AUC**: area under the curve.
[c]**ACC**: overall accuracy.
[d]**SPC**: overall specificity.
[e]**TPR**: overall true positive rate.
[f]**FNR**: overall false negative rate.

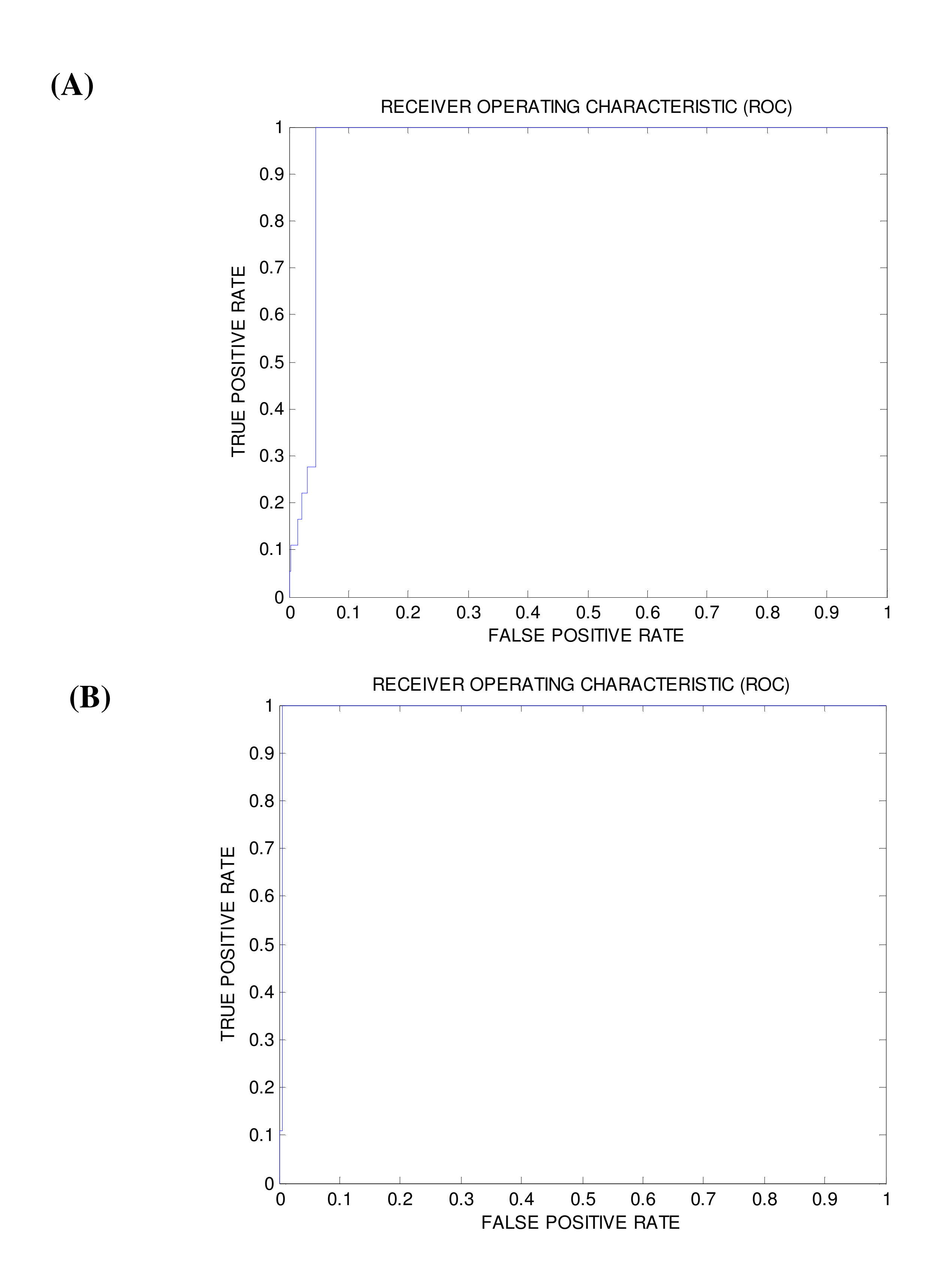


Figure 2.5: ROC curves of: **(A)** Hypo4/15, **(B)** Hypo6/35**..**

2.4.5. Comparison of pharmacophore model with the active site of ROCKII

To further emphasize the validity of our pharmacophore/QSAR modeling approach, we compared the pharmacophoric features of Hypo4/15 and Hypo6/35 and how they map training compounds **98** and **107**, respectively, with optimal docked poses of the two compounds. The docking experiments were conducted employing LigandFit docking engine and through default docking settings (CERIUS2 4.10 LigandFit User Manual; 2000). Figures 2.3 and 2.4 show the pharmacophores, docked poses and corresponding mapped conformers.

We compared the structure of **98** (IC50 = 0.02 μM, table 2.1) docked into ROCK II (PDB code: 2H9V) with Hypo4/15. Figure 2.3 shows the chemical structure of the ligand and compares its ROCK II docking picture with the way it maps Hypo4/15. Mapping the NH of the indazole ring in **98** against HBD feature in Hypo4/15 (Figure 2.3C) corresponds to hydrogen bonding interactions tying the indazole NH and the peptidic carbonyl of Ala102 (Figure 2.3A) in the docked complex. Similarly, mapping the benzene ring of the ligand's indazole against Hbic feature in Hypo4/15 (figure 2.3C) corresponds to hydrophobic stacking against the aromatic ring of Phe103. Moreover, fitting the pyrrolidine core in **98** against PosIon and Hbic features in Hypo4/15 (figure 2.3C) seems to encode for electrostatic and hydrophobic attractive interactions connecting the pyrrolidine NH and carbon ring with the carboxylate and CH_2 linker of Asp232, respectively (figure 2.3A). Finally, fitting the chlorobenzene in **98** against a RingArom feature in Hypo4/15 correlates with stacking against the adjacent hydrophobic side chain of Val106.

Similar analogies can be drawn between the docked conformer of **107** (IC_{50} = 0.009 μM, table 2.1) and the way it maps Hypo6/35, as shown in figure 2.4. Mapping the

amino functionality of the piperidine core of **107** against PosIon feature in Hypo6/35 (figure 2.4C) agrees with electrostatic attraction connecting this group and the carboxylate of Asp176 in the docked pose (figure 2.4A). Likewise, positioning the aromatic methyl substituent of **107** within a hydrophobic pocket comprised of the side chains of Leu221, Ile98, Va153, Phe384 and Met169 in the docked pose (figure 2.4A) agrees with mapping this moiety against Hbic feature in Hypo6/35 (figure 2.4C). Similarly, mapping the sp^2 nitrogen and benzene ring of the ligand's indazole group against HBA and HbicArom in Hypo6/35 (figure 2.4C) points to hydrogen bonding and aromatic stacking interactions connecting the indazole ring with the carboxylic acid and aromatic side chains of Asp232 and Phe103 in the docked pose, respectively, as in figure 2.4A.

Clearly from the above discussion, Hypo4/15 and Hypo6/35 represent two valid binding modes assumed by ligands within ROCK II. Furthermore, these models point to limited number of critical interactions required for high ligand- ROCK II affinity in each of the binding modes. In contrast, docked complexes reveal many bonding interactions without highlighting critical ones. Figures 2.3A and 2.4A only show interactions corresponding to pharmacophoric features while other binding interactions were hidden for clarity.

2.4.6. *In-Silico* Screening of Databases

Hypo4/15 and Hypo6/35 were employed as 3D search queries against the NCI (238,819 structures) using the "Best Flexible Database Search" option implemented within CATALYST (CATALYST 4.11 users' manual, 2005). Compounds that have their chemical groups spatially overlap (map) with corresponding features of the particular pharmacophoric model were captured (hits). Table 2.7 summarizes the numbers of captured hits by each pharmacophore.

NCI hits were filtered based on Lipinski's (Lipinski *et.al.*, 2001) and Veber's rules (Veber *et.al.*, 2002). Enforcing some kind of drug-likeness pre-filters should help in finding hits more amenable for subsequent optimization into leads.

Table 2.7 Numbers of captured hits by Hypo4/15, Hypo6/35.

		Pharmacophore models	
3D Database[a]	**Post screening filtering[b]**	**Hypo4/15**	**Hypo6/35**
NCI	**Before**	**674**	**4692**
	After	**172**	**215**

[a]NCI: national cancer institute list of available compounds (238,819 structures)..[b]Using Lipinski's and Veber's rules. A maximum of two Lipinski's violations were tolerated.

Surviving hits were fitted against Hypo4/15 and Hypo6/35 and their fit values, together with other relevant molecular descriptors, were substituted in QSAR equation (2.10) to predict their anti-ROCK II IC_{50} values. The highest-ranking hits compounds were evaluated *in vitro* against ROCK II assay kit (Cyclex, Japan). Figure 2.7 illustrates the dose/response plots of the active hits. To validate our bioassay settings, we determined the IC_{50} value of standard inhibitor Y-27632 under the same conditions.

Initially, hits were screened at 10 μM concentrations, subsequently; compounds showing anti-ROCK II inhibitory percentages ≥ 50% at 10 μM were further assessed to determine their IC_{50} values. Table 2.8 show active hits and their corresponding estimated and experimental anti-ROCK II bioactivities.

It remains to be mentioned that although QSAR prediction was rather accurate with one hit, i.e., **144**, it deviated significantly from experimental values with other hits (Table 2.8 and figure 2.6). We believe these errors are because training compounds used in QSAR and pharmacophore modeling are significantly structurally different from hit molecules (see section **2.4.7 Similarity Analysis between Training Compounds and Active Hits** below), which limits the extrapolatory potential of the QSAR equation. Furthermore, the fact that we implemented a bioassay method different from that used for training compounds can also explain part of the predicted-to-experimental differences in hit bioactivities. Additionally, QSAR-based predictions have their errors; in fact an r^2_{PRESS} of 0.535 suggests certain level of uncertainty in predictions.

Table 2.8. High-ranking hit molecules with their fit values against (Hypo4/15; Hypo 6/35.) their corresponding QSAR estimates from equation (2.10) and their *in vitro* anti- ROCK II bioactivities

No.[a]	Name	Structure	Fit values[b]		QSAR Predictions[c]		Experimental[d]
			Hypo4/15	Hypo6/35	log $(1/IC_{50})$	IC_{50} (µM)	IC_{50}(µM)
139	NCI0007823		0	8.12	3.405	0.0003	5.57
140	NCI0032983		0	8.25	1.12	0.076	1.41
141	NCI0104542		4.54	3.51	0.989	0.103	1.69
142	NCI0151684		0	7.81	0.625	0.237	45.35
143	NCI0380794		0	7.4	2.82	0.0015	0.70

[a]Compound numbers.
[b]Best-fit values calculated by equation (2.4).
[c]Predictions based on optimal QSAR model (2.10)
[d]Experimental *in vitro* validation against rho-kinase.

Table 2.8. High-ranking hit molecules with their fit values against (Hypo4/15; Hypo 6/35.) their corresponding QSAR estimates from equation (2.10) and their *in vitro* anti- ROCK II bioactivities

No.[a]	Name	Structure	Fit values[b]		QSAR Predictions[c]		Experimer
			Hypo4/15	Hypo6/35	log (1/IC_{50})	IC_{50} (μM)	IC_{50} (μM)
144	NCI0377095		6.45	0	-0.731	5.38	4.77
145	NCI0401383		0	7.88	2.19	0.006	1.00
146	NCI0674004		0	8.1	2.74	0.0017	1.65

[a]Compound numbers.
[b]Best-fit values calculated by equation (2.4).
[c]Predictions based on optimal QSAR model (2.10)
[d]Experimental *in vitro* validation against rho-kinase.

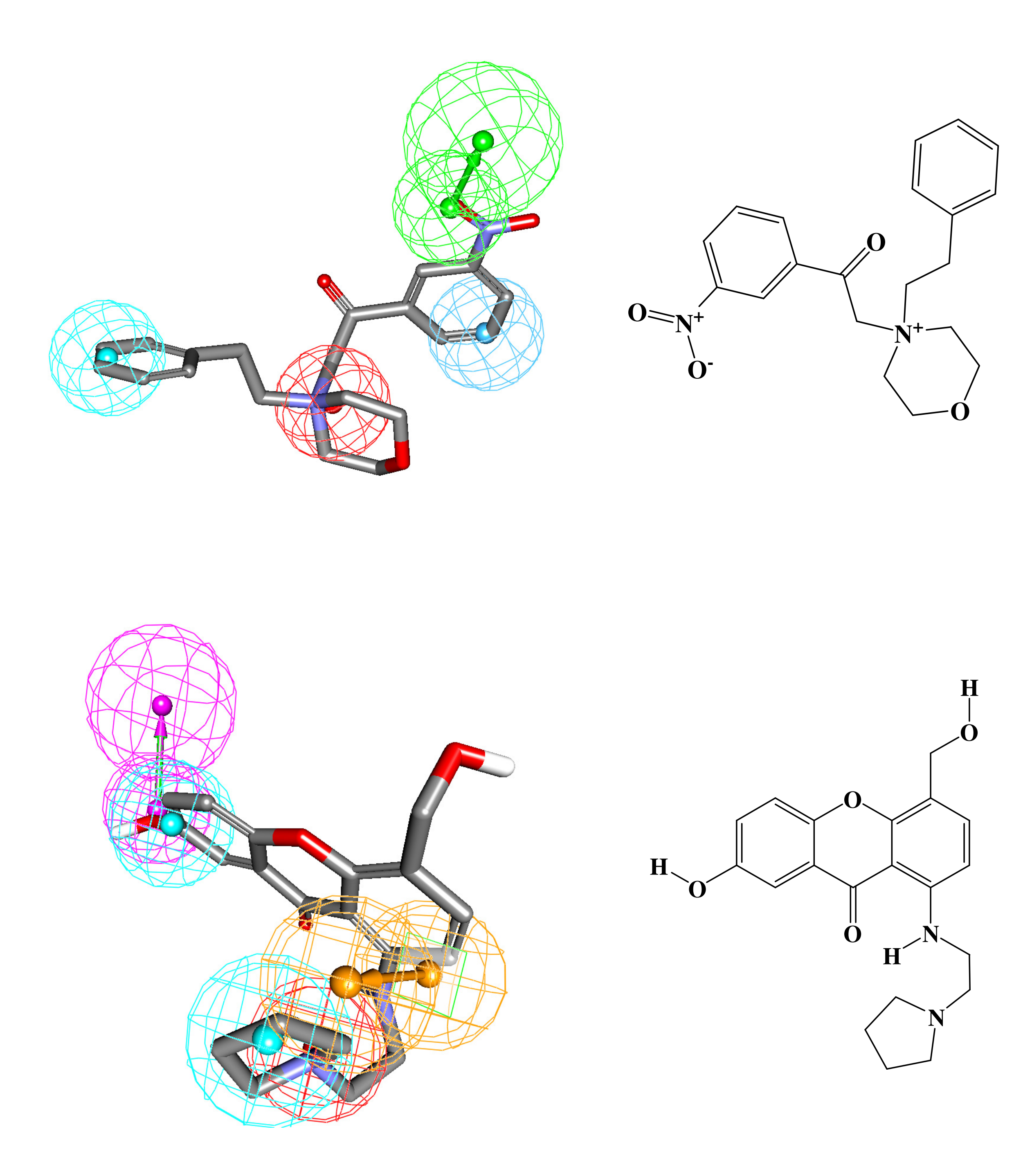


Figure 2.6: **(A)** show Hypo6/35 fitted against hit 145 (IC_{50} = 1 uM, table 2.8) **(B)** Chemical structure of 145 **(C)** show Hypo4/15 fitted against hit 144 (IC_{50} = 4.77 uM, table 2.8) **(D)** Chemical structure of 144

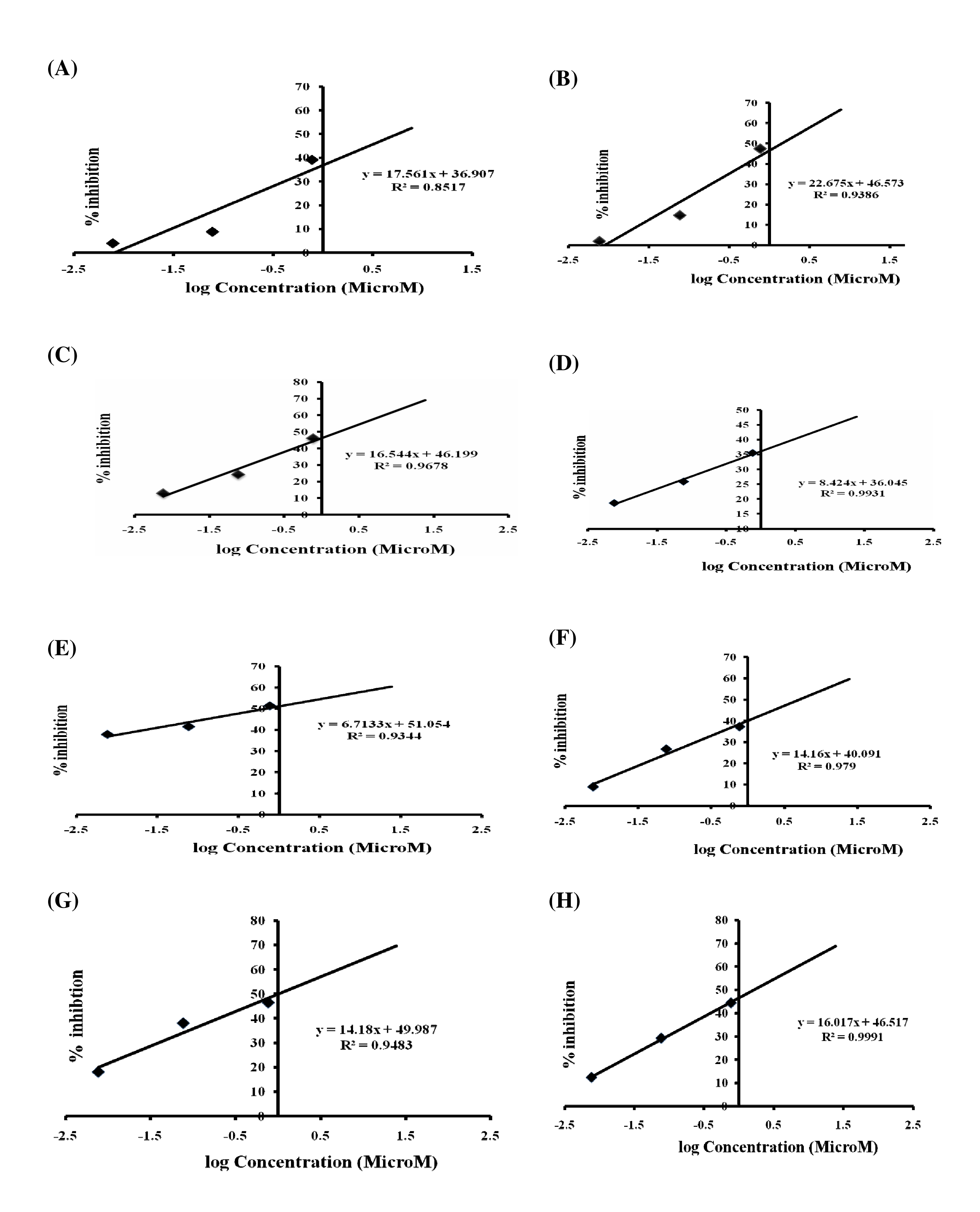


Figure 2.7: Dose-Inhibition plots for 139 (A); 140 (B); 141 (C); 142 (D); 143 (E), 144 (F), 145 (G); and 146 (H).

2.4.7 Similarity Analysis between Training Compounds and Active Hits:

To explain the significant deviation between experimental and QSAR-based predicted bioactivities of hit compounds, we decided to assess the structural similarity/diversity between training compounds (**1-138**, table 2.1, Takami *et al*, 2004, Iwakubo *et al*, 2007, a&b) (library A) and active hits (**139-146**, library B). We employed three library comparison methods implemented in Discovery Studio 2.5 to assess the structural similarity/diversity between the two groups, namely, Murcko Assemblies, Bayesian Model and Global Fingerprints.

In Murcko Assemblies, the algorithm breaks the ligands of each library into unique occurrences of molecular rings, ring assemblies, bridge assemblies, chains, Murcko assemblies, or any combination of these. Murcko assemblies are contiguous ring systems plus chains that link two or more rings (Bemis and Murcko., 1996). The two libraries are compared using a Tanimoto similarity of the assemblies based on the fragments that are common and unique to each library (Discovery Studio 2.5.5 User Guide., 2010). On the other hand, in Bayesian Model approach, two Bayesian models were built, one to learn library A and one to learn library B. Finally, it scores all ligands using both models. A distance is computed as equation 2.11:

$$\text{Distance} = \text{ScoreAA} + \text{ScoreBB} - \text{ScoreAB} - \text{ScoreBA} \quad \text{......... (2.11)}$$

where ScoreAA is the average score of library A molecules scored by the Bayesian model that learned library A molecules, while ScoreBB is the average score of library B molecules scored by the Bayesian model that learned library B. ScoreAB and ScoreBA are the average scores of libraries A and B molecules scored by the

Bayesian models that learned libraries B and A, respectively. The higher the distance, the more dissimilar the libraries are (Discovery Studio 2.5.5 User Guide., 2010).

Finally, the Global Fingerprint comparison algorithm generates a global fingerprint for all ligands in the training list and all ligands in the hits list and then computes a Tanimoto similarity coefficient between the two libraries (Discovery Studio 2.5.5 User Guide., 2010).

Table 2.9 shows the results of the three similarity/diversity assessment procedures. Clearly, the three methods suggest minimal structural similarity between known ROCK II inhibitors and our active hits, which explains the significant differences between the predicted and experimental bioactivities of active hits.

Table 2.9: Results of Similarity Analysis between Training Compounds and Active Hits.

Murcko Assemblies[a]		**Bayesian Model**[b]		**Global Fingerprints**[b,c]	
Number of total assemblies	69	Average LibA score of library A ligands	1.505	Number of total global fingerprint bits	981
Number of common	0	Average LibB score of library A ligands	-44.073	Number of common global fingerprint bits	85
Number of assemblies only	61	Average LibA score of library B ligands	-31.888	Number of global fingerprint bits only in	614
Number of assemblies only	8	Average LibB score of library B ligands	32.9765	Number of global fingerprint bits only in	282
Similarity score between the two	0	Bayesian distance between the two	110.4436	Similarity score between the two libraries	0.0866

[a]See section 1.4.7 and (Bemis and Murcko., 1996)

[b]See section 1.4.7 and (Discovery Studio 2.5.5 User Guide., 2010).

[c]Done by implementing the finger print descriptor FCFC_6 which correspond to functional-class extended-connectivity fingerprint count up to diameter 6 (Discovery Studio 2.5.5 User Guide., 2010). .

[d]Library A: The list includes all training and testing compounds employed in pharmacophore and QSAR modeling (**1-138**,table 2.1)

[e]Library B: Active hits (**139-146**, table 2.8).

2.5 Conclusion

ROCK II inhibitors are currently considered as potential treatments for hypertension. The pharmacophoric space of ROCK II inhibitors was explored via six diverse sets of inhibitors and using CATALYST-HYPOGEN to identify high quality binding model(s). Subsequently, genetic algorithm and multiple linear regression analysis were employed to access optimal QSAR model capable of explaining anti-ROCK II bioactivity variation across 138 collected ROCK II inhibitors. Two orthogonal pharmacophoric models emerged in the QSAR equation suggesting the existence of at least two distinct binding modes accessible to ligands within ROCK II binding pocket. The QSAR equation and the associated pharmacophoric models were experimentally validated by the identification of several ROCK II inhibitors retrieved via *in silico* screening, out of which 8 inhibitors illustrated micromolar potencies. Our results suggest that the combination of pharmacophoric exploration and QSAR analyses can be useful tool for finding new diverse ROCK II inhibitors.

3. Future Work

1-Dual activity testing for captured hits as Ca^{2+}/Calmodulin-Dependent Protein Kinase II (CaMKII) and Rho-Kinase inhibitors

2- Further synthesis of new s-triazine derivatives.

3- Animal studies for the activity of the potent captured hits and synthesized compounds.

5- Toxicological studies of the potent compounds

6- Exploration of the s-triazine derivatives using other targets rather than anti-hypertensive agents.

4. Appendix

Table 4.1: Values of QSAR descriptors in Equation 2.10 calculated for total Rho-kinase inhibitores in table 2.1

compound	Activ	Hypo4/15	Hypo6/35	LUMO	AtypeC24	AtypeH53	AtypeN72	CHI-0	Wiener	JX
1	0.2	0	0	-0.69	4	0	2	13.8	758	2.01
2	0.9	0	0	-0.39	5	0	2	13	653	1.99
3	35	0	0	-0.43	5	0	2	13	673	1.94
4	35	0	0	-0.09	6	0	2	12.1	572	1.93
5	35	0	0	0.09	7	0	2	11.2	495	1.89
6	2	0	0	-0.38	5	0	2	13	653	1.99
7	13.6	0	0	-0.77	5	0	2	15.5	956	2.1
8	0.8	0	0	-0.47	5	0	2	13.7	796	1.86
9	4.1	0	0	-0.67	5	0	2	13.7	818	1.82
10	7.2	0	0	-0.05	6	0	2	12.8	702	1.81
11	10.2	0	0	0.18	7	0	2	11.9	612	1.77
12	7	0	0	-0.64	5	0	2	13.7	796	1.85
13	2.2	0	0	-0.96	5	0	2	16.2	1,141	1.97
14	0.5	0	0	-0.81	4	0	1	14.5	916	1.87
15	0.9	0	0	-0.75	5	0	1	13.7	796	1.85
16	35	0	0	-0.36	4	0	2	13.8	746	2.03
17	35	0	0	-0.54	3	0	2	14.7	843	2.06
18	35	0	0	-0.52	5	0	2	14.7	867	2.02
19	35	0	0	-0.43	5	0	2	13.8	758	2
20	35	0	0	-0.34	5	0	2	13.8	758	2
21	35	0	0	-0.45	5	0	2	14.7	803	2.19
22	35	0	0	-0.39	6	0	1	13.5	848	1.7
23	35	0	0	-0.5	5	0	1	13.7	794	1.85
24	35	0	0	-0.6	5	0	1	14.5	916	1.86
25	0.5	0	0	-0.15	3	0	1	12.4	644	1.68
26	0.9	0	0	-0.17	3	0	1	13.1	783	1.57
27	0.5	0	0	-0.74	5	0	1	16.4	1,347	1.57
28	1	0	0	-0.31	6	0	2	14.8	999	1.69
29	0.3	0	0	-0.12	8	0	2	13.8	941	1.55
30	0.3	0	0	-0.06	3	4	0	10.8	459	1.72
31	0.7	0	0	-0.04	3	4	0	12.4	664	1.61
32	0.2	3.7	4.4	-0.05	8	4	0	15.5	1,399	1.29
33	35	6.7	0	0.31	9	4	0	15.5	1,399	1.3
34	0.4	5.9	4.5	-0.34	7	4	0	16.4	1,535	1.32
35	35	6	3.9	-0.11	7	4	0	17.1	1,695	1.34
36	35	3.3	5.2	-0.47	8	4	1	17.9	1,800	1.42
37	35	4.5	5.5	-0.11	8	4	0	17.1	1,657	1.38
38	35	6.1	5	-0.02	8	4	0	16.4	1,561	1.31
39	35	0	5.1	-0.01	8	4	0	16.4	1,575	1.3
40	0.8	0	4.3	-0.01	3	0	0	13.1	763	1.61
41	0.02	8.8	5.2	0.01	8	0	0	15.5	1,329	1.37
42	35	0	0	-0.07	3	2	0	10.1	378	1.78
43	0.03	7.7	5	-0.06	8	2	0	14.8	1,205	1.35
44	35	0	5.4	0.02	8	0	0	16.4	1,498	1.38
45	35	0	0	-0.16	3	0	1	12.4	644	1.68
46	35	0	0	-0.18	3	0	1	13.1	783	1.57
47	35	0	0	-0.79	5	0	1	16.4	1,347	1.57

Table 4.1: Values of QSAR descriptors in Equation 2.10 calculated for total Rho-kinase inhibitores in table 2.1

compound	Activ	Hypo4/15	Hypo6/35	LUMO	AtypeC24	AtypeH53	AtypeN72	CHI-0	Wiener	JX
48	35	0	0	-0.34	6	0	2	14.8	999	1.69
49	35	0	0	0.06	3	4	0	12.4	664	1.61
50	0.1	0	5.3	-0.62	9	4	0	16.2	1,522	1.35
51	0.3	0	4.5	-0.61	4	4	0	13.8	876	1.65
52	0.2	0	0	-0.59	4	0	0	13.1	738	1.72
53	0.2	0	5.3	-0.55	9	0	0	16.2	1,445	1.43
54	0.1	0	7.9	-0.58	4	0	0	16.1	1,328	1.57
55	0.21	0	2.8	-0.61	4	4	0	14.7	1,016	1.61
56	0.27	0	6.3	-0.58	4	0	0	15.4	1,139	1.63
57	35	0	5.3	-0.62	9	0	0	16.9	1,772	1.28
58	35	0	5	-0.59	9	1	0	16.9	1,688	1.35
59	35	0	0	-0.92	8	0	2	16.2	1,332	1.7
60	35	0	0	-0.8	7	0	2	15.5	1,095	1.78
61	35	0	0	-0.85	6	0	2	16.4	1,244	1.78
62	35	0	0	-0.74	7	0	2	16.2	1,298	1.66
63	35	0	0	-0.75	7	0	2	16.2	1,298	1.66
64	0.18	0	4.2	-0.58	4	0	0	14.5	1,034	1.59
65	0.18	0	4.5	-0.59	4	0	0	14.5	1,034	1.59
66	0.195	0	0	-0.59	4	0	0	14.1	1,015	1.39
67	0.07	0	0	-0.57	4	0	0	14.1	1,015	1.39
68	0.44	0	7.2	-0.59	9	0	0	17.6	2,031	1.24
69	0.58	0	7.6	-0.59	9	0	0	17.6	2,031	1.24
70	0.4	0	0	-1.24	8	0	2	15.5	1,202	1.58
71	0.3	0	0	-1.21	7	0	2	16.4	1,347	1.59
72	0.1	0	0	-1.22	6	0	2	17.3	1,496	1.6
73	0.1	0	0	-1.48	5	0	2	18.1	1,628	1.65
74	35	0	0	-1.38	6	0	1	17.3	1,496	1.59
75	35	0	0	-1.31	6	0	1	17.3	1,496	1.59
76	8.3	0	0	-1.23	3	0	2	13.1	656	2.04
77	0.9	0	0	-1.3	8	0	2	14.8	1,015	1.68
78	1.9	0	0	-1.27	6	0	2	16.6	1,273	1.7
79	35	0	0	-1.22	3	0	2	15.7	1,172	1.54
80	35	4.3	1	-1.26	8	4	2	19.5	2,471	1.23
81	0.02	0	0	-0.55	6	0	2	15.5	1,189	1.58
82	0.085	0	0	-0.24	6	0	2	15.5	1,189	1.58
83	0.22	0	0	-0.61	6	0	2	15.5	1,217	1.56
84	7.9	1.3	5.2	0.02	8	0	0	14.5	1,125	1.45
85	35	6	5.1	-0.09	7	0	0	17.1	1,847	1.21
86	0.136	0	4.2	-0.2	8	0	1	17.1	1,761	1.28
87	0.54	5.8	1.4	-0.38	8	4	1	17.1	1,765	1.27
88	0.018	7.2	4.7	-0.18	8	0	1	17.1	1,677	1.35
89	35	3.9	0	-0.15	8	4	2	17.8	2,034	1.21
90	0.1	4.7	4.9	-0.2	8	4	0	15.5	1,399	1.28
91	0.01	8.1	5.1	-0.15	8	0	0	15.5	1,329	1.36
92	35	3.3	5.2	-0.47	8	4	1	17.9	1,800	1.42
93	0.15	8.3	5	0.01	8	0	0	16.4	1,484	1.39
94	0.44	3.7	5.9	-0.06	7	4	0	16.4	1,559	1.3
95	0.29	6.8	8	-0.07	7	4	0	16.4	1,576	1.29
96	0.29	7	7.6	-0.09	7	4	0	16.4	1,593	1.28

Table 4.1: Values of QSAR descriptors in Equation 2.10 calculated for total Rho-kinase inhibitores in table 2.1

compound	Activ	Hypo4/15	Hypo6/35	LUMO	AtypeC24	AtypeH53	AtypeN72	CHI-0	Wiener	JX
97	0.045	8.4	6.3	-0.04	7	2	0	15.6	1,348	1.35
98	0.02	8.1	8	-0.06	7	2	0	15.6	1,364	1.34
99	0.024	8	8	-0.09	7	2	0	15.6	1,380	1.33
100	0.013	8.6	5.8	0.01	7	0	0	16.4	1,479	1.38
101	0.019	8.7	8	-0.06	7	0	0	16.4	1,496	1.37
102	0.011	0	7.8	-0.06	7	0	0	16.4	1,513	1.36
103	0.78	0	4.2	-0.01	3	0	0	13.1	763	1.61
104	0.34	8.9	6.2	-0.01	3	0	0	14.7	1,041	1.54
105	0.025	8.8	5.6	-0.06	7	0	0	16.4	1,513	1.36
106	0.01	9	5.9	-0.21	6	0	0	17.2	1,683	1.36
107	0.009	8.9	8	-0.02	7	0	0	16.4	1,513	1.36
108	0.016	8.7	7.4	-0.03	7	0	0	17.1	1,721	1.34
109	0.08	8	8	-0.07	6	0	0	18.6	2,065	1.37
110	0.085	8.5	5.2	-0.06	7	0	0	16.4	1,479	1.38
111	0.035	8.8	5.1	0.01	7	0	0	16.4	1,496	1.37
112	0.01	8.5	4.6	0.01	7	0	0	16.4	1,513	1.36
113	0.003	9.1	8.7	-1.58	7	0	0	17.9	1,880	1.36
114	0.087	8.2	0	-1.59	7	0	0	17.9	1,931	1.33
115	0.24	8.3	5.7	0.01	6	0	0	15.5	1,329	1.36
116	0.072	8.1	5.3	-0.02	5	0	0	15.5	1,329	1.36
117	0.19	8.7	3.1	-0.03	5	0	0	15.5	1,329	1.36
118	0.011	8.9	4.4	0.02	5	0	0	14.8	1,165	1.37
119	0.005	8.7	4.6	0.01	5	0	0	14.8	1,165	1.37
120	0.003	8.4	4.8	0.01	4	0	0	1F4.8	1,165	1.37
121	0.065	8.6	4.6	0	5	0	0	14.8	1,165	1.36
122	0.18	0.9	3.9	0	4	0	0	14.8	1,165	1.36
123	0.32	2.1	5	0.03	8	0	0	16.2	1,627	1.24
124	0.52	0	4.8	-0.04	8	0	0	16.2	1,627	1.24
125	0.21	0	6.1	-0.56	8	0	0	17.1	1,604	1.44
126	0.265	0	8.3	-0.63	8	0	0	17.1	1,622	1.43
127	0.185	0	8.2	-0.63	8	0	0	17.1	1,640	1.42
128	0.1	0	6.2	-0.6	8	2	0	16.4	1,467	1.41
129	0.065	0	8.3	-0.63	8	2	0	16.4	1,484	1.4
130	0.065	0	8.1	-0.65	8	2	0	16.4	1,501	1.39
131	0.095	0	6	-0.65	8	2	0	16.4	1,501	1.39
132	0.075	0	5.5	-0.59	7	2	0	17.2	1,623	1.42
133	0.205	3.7	5.6	-0.6	8	2	0	16.4	1,467	1.41
134	0.275	3.4	5.5	-0.57	8	2	0	16.4	1,484	1.4
135	0.235	0	5.1	-0.57	8	2	0	16.4	1,501	1.39
136	0.18	0	0	-2.24	4	2	1	14	746	1.97
137	35	0	0	-0.75	8	0	0	14.9	982	1.9
138	9	8.4	0	-0.4	7	0	3	15.4	1,044	1.98

Table 4.2: Values of QSAR descriptors in Equation 1.10 calculated for total CaMKδII inhibitores in table 1.1:

Cpd.	Activ	Hypo8/31	Hypo9/47	Hypo7/39	Jurs-FPSA-1	AtypeC_28	S_aaCH	Jurs-RPSA	S_dO
1	0.38	9.37	7.12	0	0.625	0	19.44	0.256	25.66
2	3.81	0	6.4	9.28	0.674	0	19.19	0.261	37.97
3	0.36	9.16	7.83	0	0.659	1	19.84	0.3	25.7
4	210	0	0	0	0.576	0	18.75	0.3	24.53
5	0.29	0	2.94	0	0.601	1	19.87	0.282	11.32
6	0.5	7.26	6.89	0	0.625	0	23.75	0.215	25.69
7	0.63	9.68	6.95	0	0.629	1	20.95	0.213	25.83
8	210	7.62	7.21	0	0.614	2	21.24	0.227	25.62
9	1.35	5.94	0	0	0.639	0	19.17	0.237	25.09
10	1.58	7.45	2.09	0	0.606	1	16.7	0.252	25.02
11	0.54	7.28	1.92	0	0.589	0	16.57	0.247	25.04
12	1.6	6.32	0	0	0.616	0	16.51	0.243	25.05
13	1.49	6.09	0.08	0	0.628	0	14.59	0.273	24.81
14	210	0	4.42	3.17	0.688	1	15.31	0.313	37.13
15	1.19	6.19	0	0	0.652	1	15.53	0.323	24.92
16	0.28	8.29	7.64	0	0.669	0	17.56	0.227	25.28
17	1.32	7.18	7.05	0	0.623	0	27.67	0.203	25.89
18	1.68	8.51	7.21	0	0.625	0	26.49	0.197	25.86
19	4.56	9.29	7.29	0	0.626	0	21.47	0.239	25.65
20	0.75	9.29	5.62	0	0.639	0	17.74	0.247	25.93
21	7.61	0	5.29	6.71	0.628	0	17.51	0.255	38.26
22	0.034	9.44	9.2	0	0.656	0	17.49	0.242	25.79
23	2.32	0	7.61	0	0.688	1	19.97	0.211	26.43
24	15.02	0	0	0	0.663	0	19.31	0.15	25.82
25	210	0.06	9.07	0	0.707	0	17.6	0.17	26.53
26	1.19	0	4.92	0	0.584	0	19.03	0.335	25.11
27	0.87	0	8.11	0	0.64	0	17.37	0.241	25.44
28	210	0	0	0	0.607	0	19.06	0.231	24.74
29	210	0	6.98	9.35	0.688	1	19.59	0.295	38
30	0.18	8.51	7.78	0	0.66	0	17.27	0.231	25.26
31	3.53	0	5.08	0	0.652	0	16.86	0.28	24.7
32	7.54	0	8.94	9.48	0.702	0	17.25	0.251	38.12
33	210	0	3.61	2.77	0.678	0	18.91	0.253	37.32
34	2.02	9.4	5.59	0	0.685	0	19.84	0.205	26.06
35	1.45	8.8	6.05	0	0.682	0	19.86	0.208	26.03
36	1.89	9.59	5.62	0	0.687	0	19.88	0.204	26.05
37	0.31	8.83	7.01	0	0.621	0	19.74	0.217	25.87
38	0.64	9.17	6.04	0	0.657	0	20.18	0.19	26.18
39	1.87	9.03	5.97	4.16	0.655	0	30.94	0.148	26.59
40	0.45	6.37	0.44	0	0.612	0	18.54	0.295	25.11
41	0.74	8.6	5.32	0	0.637	0	17.58	0.247	25.93
42	0.38	9.19	7.76	0	0.625	1	17.19	0.252	25.6
43	0.19	8.9	7.13	0	0.637	0	19.43	0.257	25.64
44	0.74	8.94	5.82	0	0.673	0	19.67	0.212	25.83
45	6	0	6.48	3.33	0.615	0	19.18	0.32	25.61
46	0.11	10.51	8.98	0	0.673	0	18	0.231	25.94
47	0.15	8.97	6.26	0	0.643	0	19.51	0.267	25.66
29	210	0	6.98	9.35	0.688	1	19.59	0.295	38
30	0.18	8.51	7.78	0	0.66	0	17.27	0.231	25.26
31	3.53	0	5.08	0	0.652	0	16.86	0.28	24.7
32	7.54	0	8.94	9.48	0.702	0	17.25	0.251	38.12
33	210	0	3.61	2.77	0.678	0	18.91	0.253	37.32
34	2.02	9.4	5.59	0	0.685	0	19.84	0.205	26.06
35	1.45	8.8	6.05	0	0.682	0	19.86	0.208	26.03
36	1.89	9.59	5.62	0	0.687	0	19.88	0.204	26.05
37	0.31	8.83	7.01	0	0.621	0	19.74	0.217	25.87
38	0.64	9.17	6.04	0	0.657	0	20.18	0.19	26.18
39	1.87	9.03	5.97	4.16	0.655	0	30.94	0.148	26.59
40	0.45	6.37	0.44	0	0.612	0	18.54	0.295	25.11
41	0.74	8.6	5.32	0	0.637	0	17.58	0.247	25.93
42	0.38	9.19	7.76	0	0.625	1	17.19	0.252	25.6
43	0.19	8.9	7.13	0	0.637	0	19.43	0.257	25.64
44	0.74	8.94	5.82	0	0.673	0	19.67	0.212	25.83
45	6	0	6.48	3.33	0.615	0	19.18	0.32	25.61
46	0.11	10.51	8.98	0	0.673	0	18	0.231	25.94

Table 4.2: Values of QSAR descriptors in Equation 1.10 calculated for total CaMKδII inhibitores in table 1.1:

Cpd.	Activ	Hypo8/31	Hypo9/47	Hypo7/39	Jurs-FPSA-1	AtypeC_28	S_aaCH	Jurs-RPSA	S_dO
47	0.15	8.97	6.26	0	0.643	0	19.51	0.267	25.66
48	0.56	8.42	5.37	0	0.668	0	19.59	0.249	25.76
49	2.41	5.79	5.2	0	0.637	0	19.48	0.243	25.68
50	0.68	8.99	4.69	0	0.649	0	19.86	0.242	25.96
51	0.01	9.17	9.25	0	0.535	0	16.83	0.26	25.63
52	0.013	9.29	9.25	0	0.555	0	17.06	0.229	25.75
53	0.038	8.8	7.48	0	0.574	0	15.44	0.266	25.52
54	0.032	8.73	9.18	0	0.512	0	13.83	0.256	25.64
55	0.046	8.27	6.79	4.2	0.544	0	16.51	0.266	25.69
56	0.11	8.7	7.17	0	0.63	0	19.27	0.29	25.53
47	0.15	8.97	6.26	0	0.643	0	19.51	0.267	25.66
48	0.56	8.42	5.37	0	0.668	0	19.59	0.249	25.76
49	2.41	5.79	5.2	0	0.637	0	19.48	0.243	25.68
50	0.68	8.99	4.69	0	0.649	0	19.86	0.242	25.96
51	0.01	9.17	9.25	0	0.535	0	16.83	0.26	25.63
52	0.013	9.29	9.25	0	0.555	0	17.06	0.229	25.75
53	0.038	8.8	7.48	0	0.574	0	15.44	0.266	25.52
54	0.032	8.73	9.18	0	0.512	0	13.83	0.256	25.64
55	0.046	8.27	6.79	4.2	0.544	0	16.51	0.266	25.69
56	0.11	8.7	7.17	0	0.63	0	19.27	0.29	25.53
57	0.17	9.77	9.03	1.23	0.622	0	17.05	0.272	25.76
58	0.021	9.76	7.88	4.75	0.634	0	16.49	0.279	38.05
59	0.38	8.49	6.77	5.53	0.579	0	15.93	0.383	37.31
60	0.19	8.45	6.7	5.01	0.608	0	16.38	0.389	37.58
61	0.28	8.43	6.7	5.21	0.642	0	16.71	0.32	38.19
63	1.28	0	6.38	0	0.677	1	25.37	0.155	0
64	3.14	0	6.44	5.86	0.64	0	17.37	0.241	25.44
65	0.92	0	7.27	4.54	0.741	1	27.9	0.146	0
66	2.87	0	6.39	5.91	0.65	1	27.02	0.152	0
67	1.69	0	6.73	9.5	0.704	1	26.31	0.201	11.63
68	0.21	0	6.89	9.62	0.656	1	25.64	0.288	11.04
68	11.85	9.51	7.94	0	0.667	0	16.82	0.248	38.66
69	12.99	0	6.58	8.95	0.663	1	25.61	0.293	11.23
70	3.42	0	6.69	9.77	0.69	1	26.52	0.269	10.86
71	2.22	0	6.7	5.8	0.761	1	27.92	0.14	0
72	0.72	0	6.82	9.59	0.713	1	26.58	0.211	11.74
73	2.26	0	6.42	8.91	0.624	1	27.28	0.169	0
74	0.17	0	6.96	9.36	0.659	1	25.82	0.256	11.49
75	0.066	0	1.57	0	0.611	1	14.9	0.497	10.88
76	0.39	0	0.05	0	0.546	1	16.96	0.336	10.87
77	0.064	0	2.7	0	0.584	1	19.69	0.305	10.87
78	0.042	0	4.99	9.96	0.498	1	17.19	0.293	10.89
79	210	0	2.96	0	0.649	1	20.34	0.191	11.46
80	210	0	2.71	0	0.655	1	20.56	0.284	10.71
81	210	0	2.47	0	0.669	1	20.61	0.237	11.57
82	0.29	9.6	7.57	0	0.616	0	23.65	0.244	25.87
83	0.77	0	2.87	0	0.527	1	21.29	0.142	0
84	0.66	0	4.94	0	0.591	1	18.54	0.233	0
85	0.009	11.98	5.8	0	0.667	1	19.73	0.109	0
86	0.24	0	5.65	6.54	0.638	1	19.59	0.13	0
87	0.93	0	5.11	0	0.688	1	19.86	0.091	0
88	0.012	11.72	7.94	0	0.63	1	19.1	0.16	0

Table 4.3: Experimental versus expected purity information (CHN elemental analysis) and high resolution mass spectrums of bioactive NCI hits (**89-98**, table 1.10).

cpd	% Carbon		% Hydrogen		% Nitrogen		High Resolution Mass Spectrum (*m/z*)	
	Expected	Found	Expected	Found	Expected	Found	Expected	Found
89	48.16	47.05	4.62	4.46	13.37	12.87	$[M+H]^+$ 450.10136	$[M+H]^+$ 450.10091
90	48.16	46.5	4.62	4.70	13.37	12.67	$[M+H]^+$ 450.10136	$[M+H]^+$ 450.10139
91	48.95	48.87	5.23	5.49	12.98	12.70	$[M+H]^+$ 431.16380	$[M+H]^+$ 431.16397
92	54.03	53.20	5.97	5.75	23.21	22.16	$[M+H]^+$ 422.16213	$[M+H]^+$ 422.16187
93	46.88	43.17	4.52	4.64	19.13	21.53	$[M+H]^+$ 440.14993	$[M+H]^+$ 440.14909
94	64.12	64.68	6.96	6.48	8.80	8.18	$[M+H]^+$ 319.14748	$[M+H]^+$ 319.14723
95	54.29	52.73	5.52	5.47	23.33	22.17	$[M+H]^+$ 420.14648	$[M+H]^+$ 420.14595

8. List of references

Abdula, A, M ; Abu Khalaf, R; Mubarak, M ; Taha, M. Discovery of New β-D-Galactosidase Inhibitors via Pharmacophore Modeling and QSAR Analysis Followed by In Silico Screening, **J. Comput. Chem**, 2011, 3, 463–482.

Abu Khalaf, R.; Abdula, A.; Mubarak, M.; Taha, M. Discovery of New β-D-Glucosidase Inhibitors via Pharmacophore Modeling and QSAR Analysis Followed by In Silico Screening, **J. Mol. Model**, DOI 10.1007/s00894-010-0737-1.

Abu Khalaf, R.; Abu Sheikha, G; Bustanji, Y.; Taha, M.O. Discovery of new cholesteryl ester transfer protein inhibitors via ligand-based pharmacophore modeling and QSAR analysis followed by synthetic exploration. **Eur. J. Med. Chem**. 2010, 45, 1598–1617.

Abu-Hammad, A. M.; Taha, M.O. Pharmacophore modeling, quantitative structure–activity relationship analysis, and shape-complemented in silico screening allow access to novel influenza neuraminidase inhibitors. **J. Chem. Inf. Model**. 2009, 49, 978–996.

Al-masri, I.M.; Mohammad, M. K.; Taha, M.O. Discovery of DPP IV inhibitors by pharmacophore modeling and QSAR analysis followed by in silico screening. **Chem. Med. Chem.** 2008, 3, 1763–1779.

Al-Nadaf, A.; Abu Sheikha, G.; Taha, M.O. Elaborate Ligand-Based Pharmacophore Exploration and QSAR Analysis Guide the Synthesis of Novel Pyridinium-based Potent β-Secretase Inhibitory Leads. **Bioorg. Med. Chem**. 2010, 18, 3088-115.

(a) Al-Sha'er, M , Taha, M.O..Discovery of Novel CDK1 Inhibitors by Combining Pharmacophore Modeling, QSAR Analysis and In Silico Screening followed by In vitro Bioassay. **Eur. J. Med. Chem**. 2010, 45, 4316–4330.

(b) Al-Sha'er, M , Taha, M.O. Elaborate Ligand-Based Modeling Reveal New Nanomolar Heat Shock Protein 90α Inhibitors J**. Chem. Inf. Model**. 2010, 50, 1706–1723.

Anderson, M. Calmodulin kinase signaling in heart: an intriguing candidate target for therapy of myocardial dysfunction and arrhythmias. **Pharmacol. Therapeut**. 2005, 106, 39-55.

Andrew, L. H, Colin, R. G. and Alexander, A. **Drug Discov. Today**. 2004, 9, 430-431

Beeley; Nigel, R. A.; Sage, C. GPCRs: An Update on Structural Approaches to Drug Discovery. Targets 2003, 2, 19–25.

Bemis G.W. and Murcko M.A. The Properties of Known Drugs. 1. Molecular Frameworks, **J. Med. Chem**. 1996, 39, 2887-2893.

Bersuker, I. B.; Bahçeci, S.; Boggs, J. E. In Pharmacophore Perception, Development, and Use in Drug Design; Güner O. F., Ed.; **International University Line**: La Jolla, CA, 2000, pp 457–473.
CATALYST 4.11 Users' Manual; **Accelrys Software Inc**.: San Diego, CA, 2005.

CERIUS2 4.10 LigandFit User Manual; Accelrys Inc.: San Diego, CA, 2000.; Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites, **J. Mol. Graphics Modell.** 2003, 21, 289 -307.

CERIUS2, QSAR Users' Manual, version 4.10; **Accelrys Inc.**: San Diego, CA, 2005; pp 43–88, 221–235, 237-250.

Clement, O.O.; Mehl, A. T. Pharmacophore Perception, Development, and Use in Drug Design. In IUL Biotechnology Series; Guner, O.F., Ed.; **International University Line:** La Jolla, CA, 2000; pp 71–84.

Curd, S.; Landquist, K.; Rose, L. Synthetic antimalarials. XII. Some 1,3,5-triazine derivatives. **J Chem S**oc, 1947, 154-160.

CycLex, CaM kinase II Assay Kit (Cat# CY-1173) Users' Manual. **CycLex Co.**, Ltd, Ina, Nagano, Japan, 2009

CycLex, Rho-Kinase Assay Kit (Cat# CY-1160) Users' Manual. **CycLex Co**., Ltd, Ina, Nagano, Japan, 2009

Davies, S; Reddy, H; Caivano, M; Cohen, P. Specificity and mechanism of action of some commonly used protein kinase inhibitors. **Biochem. J.** 2000, 351, 95-105.

DePristo, M. A.; de Bakker, P. I. W.; Blundell, T. L. Heterogeneity and Inaccuracy in Protein Structures Solved by X-ray Crystallography. Structure 2004, 12, 831–838.

Desai, P. S.; Desai, K. R. Synthesis of 2-(4'-nitroanilino)-4-(2'-carbamoylphenoxy)-6-(arylthioureido)-s-triazine derivatives as antibacterial agents. **Indian J Chem**, 1990, 688-709.

Desai, R. M. Simple and efficient synthetic routes to s-triazinyl dithiocarbamate derivatives: 2,4-diarylamino-6-[N-(4-ethoxyphenyl)dithiocarbamoyl]-s-triazines and 2,4-bis[N-(4-ethoxyphenyl)dithiocarbamoyl]-6-arylamino-s-triazines. **Indian J Chem**, 2004, 367-373.

Discovery Studio 2.5.5 User Guide, **Accelrys Inc**.: San Diego, CA, 2010

Discovery Studio version 2.5 (DS 2.5) User Manual, in, **Accelrys Inc**, San Diego, CA, 2009.

Dong, M , Yan, Bryan P., Liao, James K., Lam, Yat-Yin, Yip, Gabriel W.K., Yu, Cheuk-Man. (2010), Rho-kinase inhibition: a novel therapeutic target for the treatment of cardiovascular diseases. **Drug Discovery Today,** 15, 622-629.

Fischer, R. The Principle of Experimentation Illustrated by a Psycho-Physical ExpeHafner Publishing Co., 8th ed.; **Hafner Publishing:** New York, 1966; Chapter II

Fontana, P.; Amorosa, M.; Alfieri, L. Relation between chemical structure and biological activity of some symmetrical triazine derivatives. **Pubbl. univ. cattolica S.** Cuore, 1960, 77, 115-123.

Gohda, Keigo.; Hakoshima T. A molecular mechanism of P-loop pliability of Rho-kinase investigated by molecular dynamic simulation. **J. Comput. Aid. Mol. Des**. 2008, 22, 11, 789-797.

Ho, K , Beasley, J, Belanger, L, Black, D, Chan, J, Dunn, D, Hu, B, Klon, A, Kultgen, S, Ohlmeyer, M, Parlato, S, Ray, P, Pham, Q, Rong, Y, Roughton, A, Walker, T, Wright, J , Xu, K, Xu, Y, Zhang, L, Webba, M. (2009), Triazine and pyrimidine based ROCK inhibitors with efficacy in spontaneous hypertensive rat model. **Bioorg. Med. Chem. Lett**.,19, 6027–6031.

Hudmon, A.; Schulman, H. Structure–function of the multifunctional Ca2+/calmodulin-dependent protein kinase II. **Biochem. J.** 2002, 364, 593-611.

Hund, T.; Decker, K.; Kanter, E.; Mohler, P.; Boyden, P.; Schuessler, R.; Yamada, K.; Rudy, Y. Role of activated CaMKII in abnormal calcium homeostasis and I_{Na} remodeling after myocardial infarction: Insights from mathematical modeling. **J. Mol. Cell Cardiol.** 2008, 45, 420-482.

Irwin, J. J.; Shoichet, B. K. ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. **J. Chem. Inf. Comput. Sci**. 2005, 45, 177–182.

(a) Iwakubo, M , Takami, A , Okada, Y , Kawata, T , Tagami, Y , Ohashi, H ,Sato, M, Sugiyama, T , Fukushima, K , Iijima. H. (2007), Design and synthesis of Rho kinase inhibitors (II) **Bioorg. Med. Chem. Lett**.,15, 350–364.

(b) Iwakubo, M , Takami, A , Okada, Y , Kawata, T , Tagami, Y , Sato, M , Sugiyama, T , Fukushima, K , Shinichiro, Amano, T , Kaibuchib, K , Iijima. H. (2007), Design and synthesis of rho kinase inhibitors (III). **Bioorg. Med. Chem. Lett.**,15, 1022-1033.

Jacobsson, M.; Liden, P.; Stjernschantz, E.; Bostroem, H.; Norinder, U. Improving Structure-Based Virtual Screening by Multivariate Analysis of Scoring Data. **J. Med. Chem.** 2003, 46, 5781-5789.

Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection—What can we learn from earlier mistakes? **J. Comput. Aided. Mol**. 2008, 22, 213–228.

Klebe, G. Virtual Ligand Screening: Strategies, Perspectives and Limitations. **Drug Discovery Today** 2006, 11, 580–594.

Kohlhaas, M.; Zhang, T.; Seidler, T.; Zibrova, D.; Dybkova, N.; Steen, A.; Wagner, S.; Chen, L.; Heller Brown, J.; Bers, D. Increased Sarcoplasmic Reticulum Calcium Leak but Unaltered Contractility by Acute CaMKII Overexpression in Isolated Rabbit Cardiac Myocytes. **Circ. Res.** 2006, 98(2), 235-244.

Krovat, E. M.; Langer, T. Non-Peptide Angiotensin II Receptor Antagonists: Chemical Feature Based Pharmacophore Identification. **J. Med. Chem.** 2003, 46, 716-726.

Kumar, R , Singh, V ,. Baker, K. (2007), Kinase inhibitors for cardiovascular disease. **Journal of Molecular and Cellular Cardiology**, 42, 1–11.

Kurogi, Y.; Güner, O. F. Pharmacophore modeling and three dimensional database

searching for drug design using catalyst. **Curr. Med. Chem**. 2001, 8, 1035–1055.

Kurokawa, H.; Osawa, M.; Kurihara, H.; Katayama, N.; Tokumitsu, H.; Swindells, M.; Kainosh, M.; Ikura, M. Target-induced conformational adaptation of calmodulin revealed by the crystal structure of a complex with nematode Ca^{2+}/ calmodulin - dependent kinase kinase peptide**. J. Mol. Biol.** 2001, 312, 59-68.

Leftheris, K.; Ahmed, G.; Chan, R.; Dyckman, A.; Hussain, Z.; Ho, K.; Hynes, J.; Letourneau, J.; Li, W.; Lin, Sh., Metzger, A.; Moriarty, K.; Riviello, C.; Shimshock, Y.; Wen, J.; Wityak, J.; Wrobleski, S.; Wu, H.; Wu, J.; Desai, M.; Gillooly, K.; . Lin, T.; Loo, D.; McIntyre, K.; Pitt, S.; Shen, D.; Shuster, D.; Zhang, R.; Diller, D.; Doweyko, A.; Sack, J.; Baldwin, J.; Barrish, J.; Dodd, J.; Henderson, I.; Kanner, S.; Schieven, G.; Webb, M. The Discovery of Orally Active Triaminotriazine Aniline Amides as Inhibitors of p38 MAP Kinase. **J. Med. Chem.** 2004, 47, 6283-6291.

(a) Levy, D.; Wang, D.; Lu, Q.; Chen, Z.; Perumattam, J.; Xu, Y.; Liclican, A.; Higaki, J.; Dong, H.; Laney, M.; Mavunkel, B.; Dugar, S. Aryl–indolyl maleimides as inhibitors of CaMKIIδ. Part 1: SAR of the aryl region, **Bioorg. Med. Chem. Lett**., 2008, 18, 2390–2394.

(b) Levy, D.; Wang, D.; Lu, Q.; Chen; Z.; Perumattam, J.; Xu, Y.; Higaki, J.; Dong, H.; Liclican, A.; Laney, M.; Mavunkel, B.; Dugar, S. Aryl-indolyl maleimides as inhibitors of CaMKIIδ. Part 2: SAR of the amine tether. Bioorg**. Med. Chem. Lett.,** 2008, 18, 2395–2398.

Li, H.; Sutter, J.; Hoffmann, R. In Pharmacophore Perception, Development, and Use in Drug Design; Güner, O. F., Ed.; **International University Line:** La Jolla, CA, 2000; pp 173-189.

Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Adv. Drug Del. Rev**. 2001, 46, 3-26

Lu, Q.; Chen, Z.; Perumattam, J.; Wang, D.; Liang, W.; Xu, Y.; Do, S.; Bonaga, L.; Higaki, J.; Dong, H.; Liclican, A.; Sideris, S.; Laney, M.; Dugar, S.; Mavunkel, B.; Levy, D. Aryl–indolyl maleimides as inhibitors of CaMKIIδ. Part 3: Importance of the indole orientation. **Bioorg. Med. Chem. Lett**., 2008, 18, 2399–2403.

MacDonnell, S.; Weisser-Thomas, J.; Kubo, H.; Hanscome, M.; Liu, Q.; Jaleel, N.; Berretta, R.; Chen, X.; Brown, J.; Sabri, A. CaMKII Negatively Regulates Calcineurin-NFAT Signaling in Cardiac Myocytes. **Circ. Res.** 2009, 105(4), 316-325.

Maier, L , Bers, D. (2002) Calcium-Calmodulin Kinase II: Heartbeat to Heartbeat and

Beyond. **J. Mol. Cell. Cardio**., 34, 919-939

Mavunkel, B.; Xu, Y.; Goyal, B.; Lim, D.; Lu, Q.; Chen, Z.; Wang, D.; Higaki, J.; Chakraborty, I.; Liclican, A.; Sideris, S.; Laney, M.; Delling, U.; Catalano, R.; Higgins, L.; Wang, Y.; Wang, J.; Feng, Y.; Dugar, S.; Levy, D. Pyrimidine-based inhibitors of CaMKIIδ. **Bioorg. Med. Chem. Lett**., 2008, 18, 2404–2408.

Mirzoeva, S.; Koppal, T.; Petrova, T.; Lukas, T.; Watterson, D.; Van Eldik, L. Screening in a cell-based assay for inhibitors of microglial nitric oxide production reveals calmodulin-regulated protein kinases as potential drug discovery targets. **Brain Resear.** 1999, 844, 126–134.

Muller, B , Mack, H , Teusch, N. (2005), RHO KINASE, A Promising Drug Target For Neurological Disorders. **Nature Reviews,** 4, 387-399.

Offermanns, S. and Wettschureck, N. (2002), Rho/Rho-kinase mediated signaling in physiology and pathophysiology. **Journal of Molecular Medicine**, 80, 629–638.

Olson, K . (2008), Applications for ROCK kinase inhibition. **Current Opinion in Cell Biology**, 20, 242-248.

Poptodorov, K.; Luu, T.; Langer, T.; Hoffmann, R. In Methods and Principles in Medicinal Chemistry. Pharmacophores and Pharmacophores Searches; Hoffmann, R. D., Ed.; **Wiley-VCH**: Weinheim, Germany, 2006; 2, pp 17–47.

Ramsey, L. F.; Schafer, W. D. The Statistical Sleuth, 1st ed.; **Wadsworth Publishing Company: Belmont**, CA, 1997.

Sheridan, R. P.; Kearsley, S. K. Why do we need so many chemical similarity search methods? **Drug Discov. Today**, 2002, 7, 903–911.

Shimokawa, H , Rashid, M.(2007), Development of Rho-kinase inhibitors for cardiovascular medicine. **Trends in Pharmacological Sciences**, 28, 296-302.

Sossalla, S.; Fluschnik, N.; Schotola, H.; Ort, K.; Neef, S.; Schulte, T.; Wittkoepper, K.; Renner, A.; Schmitto, J.; Gummert, J. Inhibition of Elevated Ca^{2+}/Calmodulin-Dependent Protein Kinase II Improves Contractility in Human Failing Myocardium. **Circ. Res**. 2010, 107(9), 1150-1161.

Steuber, H.; Zentgraf, M.; Gerlach, C.; Sotriffer, C. A.; Heine, A.; Klebe, G. J. Expect the Unexpected or Caveat for Drug Designers: Multiple Structure Determinations Using Aldose Reductase Crystals Treated under Varying Soaking and Co-crystallisation Conditions. Mol.Biol. 2006, 363, 174–187.

Stolen, T.; Hoydal, M.; Kemi, O.; Catalucci, D.; Ceci, M.; Aasum, E.; Larsen, T.; Rolim, N.; Condorelli, G.; Smith, G. Interval Training Normalizes Cardiomyocyte Function, Diastolic Ca^{2+} Control, and SR Ca^{2+} Release Synchronicity in a Mouse

Model of Diabetic Cardiomyopathy. **Circ. Res**. 2009, 105(6), 527-536.

Stubbs, M. T.; Reyda, S.; Dullweber, F.; Moller, M.; Klebe, G.; Dorsch, D.; Mederski, W.; Wurziger, H. pH-Dependent Binding Modes Observed in Trypsin Crystals: Lessons for Structure-Based Drug Design. **ChemBioChem** 2002, 3, 246–249.

Sutter, J.; Güner, O.; Hoffmann, R.; Li, H.; Waldman, M. In Pharmacophore Perception, Development, and Use in Drug Design; Güner, O. F., Ed.; **International University Line**: La Jolla, CA, 2000; pp 501–511.

Taha, M.O.; Bustanji, Y.; Al-Bakri, A.G.; Yousef, M.; Zalloum, W.A.; Al-Masri, I.M.; Atallah, N. Discovery of new potent human protein tyrosine phosphatase inhibitors via pharmacophore and QSAR analysis followed by in silico screening. **J. Mol. Graphics Modell**. 2007, 25, 870–884.

(a) Taha, M.O.; Atallah, N.; Al-Bakri, A.G.; Paradis-Bleau, C.; Zalloum, H.; Younis, K.; Levesque, R.C. Discovery of New MurF Inhibitors via Pharmacophore Modeling and QSAR Analysis followed by in-silico screening. **Bioorg. Med. Chem**. 2008, 16, 1218-1235.

(b) Taha, M.O.; Bustanji, Y.; Al-Ghussein, M.A.S.; Mohammad, M.; Zalloum, H.; Al-Masri, I.M.; Atallah, N. Pharmacophore Modeling, Quantitative Structure-Activity Relationship Analysis and In Silico Screening Reveal Potent Glycogen Synthase Kinase-3β Inhibitory Activities for Cimetidine, Hydroxychloroquine and Gemifloxacin, **J. Med. Chem**. 2008, 51, 2062-2077.

(c)Taha, M.O.; Dahabiyeh, L. A.; Bustanji, Y.; Zalloum, H.; Saleh, S. Combining Ligand-Based Pharmacophore Modeling, QSAR Analysis and In-Silico Screening for the Discovery of New Potent Hormone Sensitive Lipase Inhibitors. **J. Med. Chem**. 2008, 51, 6478–6494.

Taha, M.O.; Trarairah, M , Zalloum, H , Abu Sheikha. G Pharmacophore and QSAR Modeling of Estrogen Receptor B Ligands and Subsequent Validation and In Silico Search for New Hits. **J Mol Graph Model.** 2010, 28, 383-400.

Takami, A , Iwakubo, M , Okada, Y , Kawata, Odai, H , Takahashi, N , Shindo, K , Kimura, K , Tagami, Y , Miyake, M , Fukushima, K , Inagaki, M , Amano, M , Kaibuchi , K , Iijima. H. (2004), Design and synthesis of Rho kinase inhibitors (I). **Bioorg. Med. Chem. Lett**.,12, 2115–2137

Tamura, M ; Nakao, H; Yoshizaki, H ; Shiratsuchi, M ; Shigyo, H ; Yamada, H ; Ozawa, T ; Totsuka, T ; Hidaka, H. Development of specific Rho-kinase inhibitors and their clinical application , **Biochim Biophys Acta** , 2005, 1754, 245 – 252

Triballeau, N.; Acher, F.; Brabet, I.; Pin, J-P.; Bertrand, H-O. Virtual Screening Workflow Development Guided by the "Receiver Operating Characteristic" Curve

Approach. Application to High-Throughput Docking on Metabotropic Glutamate Receptor Subtype 4. **J. Med. Chem.** 2005, 48, 2534–2547.

Van Drie, J.H. Pharmacophore discovery— lessons learned, **Curr. Pharm. Des**. 2003; 9: 1649–1664.

Veber, D.F.; Johnson, S.R.; Cheng, H.Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular properties that influence the oral bioavailability of drug candidates. **J. Med. Chem.** 2002, 45, 2615-2623.

(a) Yamaguchi, H. Kasa, M., Amano, M., Kaibuchi, K., Hakoshima, T. Molecular mechanism for the regulation of rho-kinase by dimerization and its inhibition by fasudil. **Structure,** 2006, 14, 589-600

(b) Yamaguchi, H., Miwa, Y., Kasa, M.,Kitano, K., Amano, M., Kaibuchi, K., Hakoshima, T. Structural basis for induced-fit binding of Rho-kinase to the inhibitor Y27632. **J.Biochem**. 2006, 140, 305-311.

Verdonk, M. L.; Marcel L.; Berdini, V.; Hartshorn, M. J.; Mooij, W.T. M.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual screening using protein-ligand docking: Avoiding artificial enrichment. J. Chem. Inf. Comput. Sci. 2004, 44, 793–806.

Wolf, N.; Schuldt, H.; Baldwin, M. s-Triazine derivatives-new class of fungicides. Science, 1955, 121, 61-62.

الاكتشاف و التحسين الكيماوي لمثبطات جديدة لرو كاينيز و\أو كالسيوم كالموديولن-ديبنديت بروتين كاينيز II كعلاجات محتملة لمرض ارتفاع ضغط الدم و ذلك بواسطة النمذجة الجزيئية المحوسبة.

إعداد
رند عمر محمد شاهين

المشرف
الأستاذ الدكتور معتصم طه الحوامدة

ملخص

تم استكشاف الفراغ الفارماكوفوري للأنزيمات بإستخدام عدة مجموعات مختلفة من المثبطات المعروفة لأنزيمي **الرو كاينيز و الكالسيوم كالموديولن-ديبنديت بروتين كاينيزII** متبوعة بعمل لوغاريتم جيني (genetic algorithm) و تحليل تمثيلي خطي (multiple linear regression) لاختيار أفضل مجموعات من النماذج المحوسبة (pharmacophoric models) مع موصفات (descriptors) فيزيائية ثنائية الأبعاد قادرة على تفسير التغير الكمي للفاعلية مقارنة مع التركيز.

ظهرت ثلاث نماذج **لإنزيم الكالسيوم كالموديولن-ديبنديت بروتين كاينيزII** ونموذجين لأنزيم **الرو كاينيز** بمعادلة ناجحة لكل منهما وكانت معاملات المعادلة الناجحة لأنزيم **الكالسيوم كالموديولن-ديبنديت بروتين كاينيزII** هي:

$r^2 = 0.70$, F-statistic = 18.19, $r^2_{LOO} = 0.71$, $r^2_{PRESS} = 0.60$

و معاملات المعادلة الناجحة **لأنزيم الرو كاينيز** هي:

r = 0.82, F-statistic = 18.75, $r^2_{LOO} = 0.68$, $r^2_{PRESS} = 0.54$

ثم تم تقييم النماذج المحوسبة الناجحة بواسطة منحنياتها المستقبلة-العاملة المميزة (ROC curves). تم تأكيد فعالية النماذج المستخلصة عمليا بالتعرف على عدة مثبطات فعالة من قواعد بيانات جزيئية محوسبة، من خلال نتائج التحليل ظهر مثبطين تصل فعاليتهما إلى مستوى 20 و 82 نومولار لأنزيم **الكالسيوم كالموديولن-ديبنديت بروتين كاينيزII**. كما تم إكتشاف عدة مثبطات لأنزيم الرو كاينيز تتميز فعاليتها بكونها منخفضة وتصل إلى حدود ميكرومولية منخفضة. تم استخدام النمذجة الجزيئية المحوسبة. ثم اخترنا النموذج الأفضل لأنزيم **الكالسيوم**

كالموديولن–ديبندينت بروتين كاينيز II بهدف تـــصنيع مركبـــات ذات فعاليـة بنسب مختلفة يصل الأكثر فعالية إلى 154 نانومولار.